Preface

This review provides an independent and evidence-based assessment of the effectiveness of United Nations Children’s Fund (UNICEF) humanitarian action and development programming. The approach and methodology to reviewing multilateral organizations was developed under the guidance of the Development Assistance Committee Network on Development Evaluation (DAC-EVALNET) of the Organisation for Economic Co-operation and Development (OECD). It complements the periodic assessments done by the Multilateral Organization Performance Assessment Network (MOPAN) of the multilateral organizations’ operational effectiveness. By focusing on development effectiveness and carefully selecting assessment criteria, the reviews seek to avoid duplication or overlap with the MOPAN process.

The review was planned in consultation with the Evaluation Office of UNICEF and led by the Policy and Operations Evaluation Department of the Netherlands Ministry of Foreign Affairs. It relies on the content of evaluation reports produced by UNICEF supplemented with a review of UNICEF corporate documents and consultation with staff at UNICEF headquarters in New York. The methodology applied is a structured meta-synthesis of the content of a purposive sample of 62 evaluations that is illustrative of UNICEF’s programming across a number of dimensions. The analysis applied a common set of assessment criteria derived from the DAC evaluation criteria.

The review was conducted by a team from the consulting firm, Goss Gilroy Inc. (Ted Freeman, Sheila Dohoo Faure, Louise Mailloux, Tasha Truant and Robert Barnes). The team is grateful to UNICEF for its helpfulness and useful, practical advice.

The content of this report does not necessarily reflect the official positions, policies or views of the OECD or the Netherlands Ministry of Foreign Affairs. The report is designed for use by UNICEF, its Governing Board, the members of the DAC-EVALNET and other interested stakeholders.

We would like to thank everyone who contributed to this review for their valued input and support.

Ted Kliest
Senior Evaluator
Policy and Operations Evaluation Department
Netherlands Ministry of Foreign Affairs

Martyn Pennington
Chair
Network on Development Evaluation
Development Assistance Committee
OECD
## Table of Contents

**Preface** ................................................................................................................. i

**Table of Contents** ................................................................................................. ii

**Executive Summary** ............................................................................................. v

**Acronyms** ............................................................................................................... xiv

### 1.0 Introduction ....................................................................................................... 1

1.1 Background ........................................................................................................ 1

1.2 Purpose ............................................................................................................... 2

1.3 Structure of the Report ...................................................................................... 2

1.4 UNICEF: A Global Organization Focused on Equity for Children .............. 3

1.5 Program Evaluation at UNICEF ....................................................................... 7

1.6 Monitoring and Reporting Results .................................................................. 11

### 2.0 Methodology .................................................................................................... 15

2.1 Evaluation Population and Sample ................................................................. 15

2.2 Criteria ............................................................................................................. 17

2.3 Review Process and Data Analysis ................................................................. 18

2.4 Other Data Collection ..................................................................................... 19

2.5 Limitations ....................................................................................................... 19

### 3.0 Findings on UNICEF’s Development Effectiveness ...................................... 21

3.1 Relevance ......................................................................................................... 23

3.2 Objectives Achievement ................................................................................ 29

3.3 Cross-Cutting Themes .................................................................................... 36

3.4 Sustainability ................................................................................................... 43

3.5 Efficiency ......................................................................................................... 49

3.6 Monitoring and Evaluation ............................................................................. 56

### 4.0 Conclusions .................................................................................................... 62
Annexes ..................................................................................................................65

Annex 1: Effectiveness Criteria .................................................................66
Annex 2: Evaluation Sample ...........................................................................67
Annex 3: Methodology ....................................................................................70
Annex 5: Global Evaluations and UNICEF Documents ...............................89
Annex 6: Selected Results for Comparisons of Least-Developed and Middle-Income Countries and by Focus Area and Humanitarian Action ..........92
Annex 7: UNICEF Staff Interviewed ...............................................................97

Figures

Figure 1: UNICEF Activities and Programs in Five Focus Areas of the MTSP ........4
Figure 2: UNICEF Expenditures, by Focus Area and Humanitarian Action, 2009-2011 ...6
Figure 3: UNICEF Expenditures, by Type, 2009 - 2011 ....................................7
Figure 4: Organizational Responsibilities for Program Evaluation at UNICEF ..8
Figure 5: Performance Indicators for the Evaluation Function at UNICEF 2011 ..9
Figure 6: Coverage and Summary Results for Each Sub-criterion ...............21
Figure 7: Number of Evaluations Addressing Sub-criteria for Relevance .......23
Figure 8: Findings for Relevance .................................................................24
Figure 9: Number of Evaluations Addressing Sub-criteria for Objectives Achievement .30
Figure 10: Findings for Objectives Achievement .........................................31
Figure 11: Number of Evaluations Addressing Sub-criteria for Cross-Cutting Themes ...36
Figure 12: Findings for Effectiveness in Supporting Gender Equality .........39
Figure 13: Number of Evaluations Addressing Sub-criteria for Sustainability ...44
Figure 14: Findings for Sustainability ...........................................................45
Figure 15: Number of Evaluations Addressing Sub-criteria for Efficiency .......50
Figure 16: Findings on Efficiency .................................................................51
Figure 17: Number of Evaluations Addressing Sub-criteria for Using Evaluation and Monitoring ...............................................................57
Review of UNICEF’s Development Effectiveness

Figure 18: Findings for Using Evaluation and Monitoring .................................58
Figure 19: Development of Populations of UNICEF Evaluations......................71
Figure 20: Quality Review Grid ..................................................................72
Figure 21: Results of Quality Review Scoring ..............................................74
Figure 22: Profile of Population and Sample of Evaluations, by Year ...........75
Figure 23: Profile of Sample of Evaluations, by Range of Years Covered ........75
Figure 24: Profile of Sample of Evaluations, by Commissioning Office .........76
Figure 25: Profile of Population and Sample of Evaluations, by Region ........76
Figure 26: Profile of Population and Sample of Evaluations, by Country Classification .76
Figure 27: Profile of Population and Sample of Development Evaluations, by Funding and Country, by Funding for 2007 – 2011 ..................................................77
Figure 28: Profile of Population and Sample of Humanitarian Evaluations, by Funding and Country, by Funding for 2007 - 2011 ..............................................................78
Figure 29: Profile of Population and Sample of Evaluations, by MTSP Focus Area and Humanitarian Action .................................................................79
Executive Summary

Background

This report presents the results of a review of the effectiveness of development and humanitarian programming supported by the United Nations Children’s Fund (UNICEF). The common approach and methodology for this review were developed under the guidance of the Organisation for Economic Co-operation and Development’s Development Assistance Committee (DAC) Network on Development Evaluation (DAC-EVALNET) and relies on the content of published evaluation reports produced by UNICEF, supplemented with a review of UNICEF corporate documents and consultation with staff at UNICEF headquarters in New York. The review was commissioned by the Ministry of Foreign Affairs of the Netherlands and carried out by a team from Goss Gilroy Inc. of Ottawa, Canada.

Purpose

The purpose of the review is to provide an independent, evidence-based assessment of the humanitarian and development effectiveness of UNICEF operations (hereafter referred to as “programs”) for use by all stakeholders. The approach is intended to work in a coordinated way with initiatives such as the DAC-EVALNET/United Nations Evaluation Group (UNEG) Peer Reviews of United Nations organization evaluation functions and assessments carried out by the Multilateral Organizations Performance and Assessment Network. It also recognizes that multilateral organizations continue to make improvements and strengthen their reporting on development effectiveness and should eventually be providing regular, evidence-based, field-tested reporting on effectiveness themselves.

Approach and Methodology

The review, carried out from October 2012 to May 2013, began with a preliminary review of UNICEF documents and the identification of the population of UNICEF evaluation reports found in UNICEF’s Global Database and reports from its Global Evaluation Report Oversight System. Interviews were then conducted with evaluation

Five UNICEF Program Focus Areas

1. Young child survival and development
2. Basic education and gender equality
3. HIV/AIDS and children
4. Child protection from violence, exploitation and abuse
5. Policy advocacy and partnerships for children’s rights

Medium-Term Strategic Plan, 2006 – 2013
and program staff at UNICEF headquarters in New York. In consultation with evaluation staff, the review team drew a sample of 70 UNICEF evaluation reports, published between 2009 and 2011, that was illustrative of UNICEF programming. Evaluations were selected to include development and humanitarian programming, programming from the various regions and types of countries in which UNICEF works and programming from each of the five focus areas in UNICEF’s Medium-Term Strategic Plan. However, it should be noted that the population of evaluations identified at UNICEF under-represented humanitarian action and programming in some of the countries receiving the largest amounts of development or humanitarian funding.

The quality of each evaluation was assessed against criteria derived from the UNEG Norms and Standards for evaluation. Since the sample did not include any evaluations found to be “not confident to act” by UNICEF’s own Global Evaluation Report Oversight System, only four reports were excluded due to quality concerns. Four others were excluded, either because they were too narrowly focused, did not provide coverage of a minimum number of effectiveness criteria or duplicated evaluation results reported in other evaluations in the sample. As a result, 62 evaluation reports were retained for systematic rating in the review.

Each evaluation report was rated, using a four-point scale that ranged from Highly Unsatisfactory to Highly Satisfactory, on six key Development Effectiveness Criteria and nineteen sub-criteria. The review team also identified the factors contributing to both positive and negative findings for each of six key criteria used to assess the development effectiveness as reported by the evaluations. The results of the meta-synthesis of evaluation findings were then summarized and presented to UNICEF staff prior to the development of this report.

### Assessment Criteria

1. Relevance of Interventions
2. The Achievement of Humanitarian and Development Objectives and Expected Results
3. Cross Cutting Themes (Environmental Sustainability and Gender Equality)
4. Sustainability of Results/Benefits
5. Efficiency
6. Using Evaluation and Monitoring to Improve Humanitarian and Development Effectiveness

### Coverage of the Assessment Criteria in Reviewed Evaluations

The review established ranges for assessing how well the sub-criteria were covered in the 62 evaluations, based on the number of evaluations that addressed each of the sub-criteria. Coverage could be strong, moderate or weak coverage. The results for sub-
criteria with weak coverage were not reported. For only two sub-criteria was the coverage rated as weak: the extent to which program-supported changes are environmentally sustainable and the effectiveness of systems for results-based-management. The findings for these sub-criteria are not reported.

**Limitation: Retrospective Nature of the Review**

The methodology section of the report describes in more detail some limitations of the review related to sampling issues and coverage. It is worth noting here that a review of development effectiveness that relies mainly on published evaluation documents is inherently retrospective, rather than forward-looking. Some important issues identified in the evaluations reviewed have been or are already being addressed by UNICEF. As one prominent example, the results of recent UNICEF global evaluations have been used by Programme Division in the preparation of inputs to next strategic plan. Where this has been pointed out to the team by UNICEF staff or where there are examples in the documents of UNICEF taking action, the actions are noted in this report.

**Findings of the Development Effectiveness Review**

**Relevance of UNICEF-Supported Programs**

The relevance of UNICEF’s programming is very well covered in the evaluations and the findings reflect that programming is highly relevant to the needs of the target groups. The vast majority of evaluations reported high suitability to target group needs, alignment with national development goals and effective partnerships. However, programming in middle-income countries was somewhat more likely to have a higher rating for effective partnerships than that in least-developed countries.

Alignment of UNICEF’s programming with government priorities was the major factor contributing to the relevance of programming, particularly to the extent that programming was designed to support the implementation of public policies, support programming and capacity building in national governments or institutions and was based on partnerships with key stakeholders. When these conditions were not met, notably when there are gaps in UNICEF’s partnerships with national government (often because government is not taking adequate ownership) and other UN agencies, this detracts from the relevance of programming.
**Objectives Achievement**

Objectives achievement is generally very well covered in the evaluations, with the exception of the sub-criterion addressing changes in national development policies and programs. The findings are also largely positive. Three-quarters of the evaluations reported positive findings on the extent to which UNICEF-supported programs achieve their stated objectives. Positive results were even stronger for the programs’ ability to provide positive benefits for the target population, with nearly nine out of ten evaluations reporting satisfactory or better findings. They were somewhat less positive on the scale of program benefits measured in terms of the number of program beneficiaries. Finally, three-quarters of the programs were rated positively for their ability to support positive changes in national policies and programs.

It was the high quality of program design (in terms of the focus, alignment and engagement of national partners) that contributed positively to objectives achievement. UNICEF’s role in influencing national policy and the appropriateness of the scope and resources for programs also contributed positively to objectives achievement. However, the opposite was also found. Weaknesses in program design, inadequate budgets and insufficient outputs detracted from the achievement of objectives.

**Gender Equality**

The evaluations presented a challenging picture of the effectiveness of gender equality. Firstly, the coverage of the effectiveness of gender equality in the evaluations was not high. Then fewer than half the evaluations that did address gender equality contained findings of satisfactory or better for this sub-criterion. Considering that 20 evaluations did not address gender equality, another way of expressing the result would be to note that only one-third of evaluations reflected that the programs effectively addressed gender equality. This finding is surprising given that addressing gender equality represents a “foundation strategy” for UNICEF programming. In addition, it seems surprising that findings of effective gender equality were somewhat more likely to be reported in evaluations of programs in least-developed countries than in middle-income countries.

UNICEF has recently been engaged in efforts to improve effectiveness in the area of gender equality. In response to the 2008 Gender Evaluation, the agency launched the Strategic Priority Action Plan for Gender Equality Action Plan 2010-2012 with eight areas of organizational action identified. UNICEF also reports its intent to take visible action on gender equality in the development of its new strategic plan.
As with other sub-criteria, the presence of specific factors contributed to effectiveness; whereas the absence of these same factors detracted from effectiveness. The factor identified most often as contributing to positive results in gender equality was the inclusion of gender-specific objectives and targets and the most common factors detracting from gender equality effectiveness were the lack of any specific gender equality objectives and/or the lack of a gender perspective in program design and implementation.

**Sustainability**

Overall the results with respect to sustainability are mixed. Generally the coverage of the sustainability sub-criteria in the evaluations was good. About half the evaluations reflected positive ratings for the likelihood that benefits would continue and the support for sustainability through developing institutional and/or community capacity. Evaluations of programming in middle income countries appear to be somewhat more likely to get positive ratings for institutional and/or community capacity building than those from least-developed countries. Three-quarters of the evaluation reports were more positive about the strengthening of the enabling environment for development. This may be explained by the fact that efforts to improve the enabling environment do not, in themselves, have to be ongoing activities and, as a result, it may be easier to achieve positive findings for this sub-criterion than for other sub-criteria, which require evidence of the long-term impacts.

Strong national government and community engagement contributed positively to sustainability. On the other hand, weaknesses in program design and implementation (notably the failure to plan for sustainability), the lack of adequate capacity at the national level and an ongoing dependence on donor support detracted from sustainability.

**Efficiency**

The coverage of efficiency in the evaluations ranged was relatively good but the findings for efficiency are mixed and require careful interpretation. The fact that two-thirds of the evaluations reported positive findings on the cost efficiency of programs represents a reasonably good result.

Nonetheless, it is a concern that, of the evaluations that addressed these sub-criteria, half reported negative findings on timeliness and two-thirds reported negative findings on the efficiency of administrative systems. Also worrying is the fact that about one in five reported a finding of highly unsatisfactory for both sub-criteria. However, since there...
may be a tendency to under-report findings with respect to timeliness and the efficiency of administrative systems when no problems are encountered at field level, the results may be biased toward negative findings.

Effective monitoring systems, the use of low-cost approaches to service delivery and strong financial planning and cash-flow management contributed to cost efficiency. On the other hand, the lack of regular and timely reporting of appropriate cost data detracted from cost efficiency. Some evaluations also noted that delays in programs start-up and the supply of inputs was a factor that reduced cost efficiency.

Similar themes are reflected in the factors that contributed to the timeliness of, and the effectiveness of systems and procedures for, program implementation and follow-up. The strong management and programming capacity of UNICEF country and regional offices, timely coordination among partner organizations, effective program monitoring and regular supervision and effective cash flow management all contributed positively to a rapid respond to program demands. The factors that detract from timeliness are related to UNICEF’s administrative systems and procedures, notably lengthy delays in procurement and rigid and cumbersome financial systems and procedures for funds disbursement delayed program implementation.

**Using Evaluation and Monitoring to Improve Humanitarian and Development Effectiveness**

Coverage for the sub-criteria regarding the use of monitoring and evaluation was strong and the results are somewhat positive. Between half and two-thirds of the evaluations received positive ratings for the effectiveness of systems and processes for evaluation, and for the use of evaluation, to improve effectiveness. However, only just over one-third of the evaluations got a positive rating for the effectiveness of systems and processes for results monitoring and reporting.

The evaluations identified few factors that contributed positively to use of monitoring and evaluation results to improve development effectiveness, beyond the fact that being part of a broader evaluation process (either UNICEF’s country program evaluations or a donor-led global evaluation) contributed to the use of monitoring and evaluation. The use of evaluation is supported by the increasing tendency to prepare evaluation management responses that include action plans for the implementation of evaluation recommendations. However, two factors detracted from the use of monitoring and evaluation: the lack of a clear results framework and inadequate or inappropriate indicators and inadequate baseline information.
Since 2011, UNICEF has engaged in a very significant effort to strengthen program results definition, monitoring and reporting, as it developed and implemented the Monitoring Results for Equity System (MoRES). This system is to be subject to a formative evaluation to be carried out by UNICEF’s Evaluation Office in 2013. In addition, UNICEF recently used existing global evaluation reports to inform the development of its upcoming strategic plan. These initiatives appear to contribute to addressing gaps in monitoring and evaluation that were identified in the current review of evaluations.

**Conclusions**

- UNICEF-supported programs are highly relevant to the needs of target group members and are supportive of the development plans and priorities of program countries. UNICEF has also had success in developing effective partnerships with government and NGOs, especially in responding to humanitarian situations. The relevance of UNICEF programming has also been supported by efforts to ensure alignment with key national development plans and policy documents.

- UNICEF has largely been effective in achieving the objectives of the development and humanitarian programs, in securing positive benefits for target group members and in supporting positive changes in national policies and programs. Where UNICEF programs have achieved success in objectives achievement it has often been supported first by high quality program design and by UNICEF’s role in influencing national sector policies. On the other hand, when UNICEF-supported programs failed to meet their objectives, it was most often due to weaknesses in project design, often linked to unclear causal relationships and lack of a results orientation.

- These weaknesses in project design were contributing factors to mixed results with respect to sustainability. UNICEF achieves fairly positive ratings for its contributions to strengthening the enabling environment for development. However, the results for the likely continuation of results or the development of institutional and community capacity are not as good. A key factor to explain these results is the failure to plan for sustainability and integrate sustainability into program designs.

- UNICEF’s performance with respect to gender equality is a serious concern. Coverage of gender equality in the evaluations was not strong and, for those that did address it, the results were weak. A major factor contributing to these poor results was the lack of specific objectives regarding gender equality or the lack of
a gender perspective in program design and implementation. Given the identification of gender equality as a foundation strategy at UNICEF, it is surprising that gender is adequately addressed in only two-thirds of the evaluations.

- There was insufficient coverage of environmental sustainability to warrant presentation of the results. It appears that environmental sustainability or the impact of UNICEF-supported programs on their environment is not addressed in most evaluations, although some evaluations of humanitarian action and programs in water, sanitation and hygiene did address the issue. Given increasing emphasis on programs to mitigate the effects of climate change in some UNICEF country programs, coverage of environmental sustainability may be expected to improve in future evaluation.

- The results for efficiency of programming are mixed and the interpretation of the results is difficult because efficiency is not covered systematically in all evaluations. It is likely that factors related specifically to timeliness and implementation systems are only addressed in evaluations if they are problematic and can help to explain weakness in objectives achievement. However, the sub-criterion related to cost efficiency was covered more systematically in the evaluations and showed somewhat positive results. However, this reflects ways in which programs have tried to reduce unit costs and increase the efficiency of resource use, rather than an analysis of overall program costs, as these costs were not identified in nearly half the evaluations. The factors that enhance cost efficiency include the establishment of effective monitoring systems to track costs and improve efficiency through lower delivery costs or unit prices and efforts to identify low-cost approaches to programming. However, where the results are not so positive, the detracting factor is most often the lack of appropriate cost data.

- The evaluations reflected somewhat positive findings with respect to the use of evaluation at UNICEF, but less so for monitoring and reporting. The use of evaluations is supported by an increasing tendency to prepare evaluation management responses that include action plans. There were insufficient ratings for the use of results-based management systems to report on the results. The lack of clear results frameworks and appropriate indicators and baseline information were factors that detracted from UNICEF’s effective use of monitoring and evaluation systems. The current development and implementation of the MoRES represents a significant effort to address this issue.

- While the sample of evaluations reviewed reflects reasonably the population of UNICEF evaluations between 2009 and 2011, a review of the profile, by country
and by type of programming, of the evaluations conducted suggests that UNICEF does not have adequate coverage of the programming in the countries that received the largest amounts of both development and humanitarian funding. There is noticeably limited coverage of UNICEF’s humanitarian action.

- UNICEF is investing considerable time and effort into the development of systems for results monitoring being implemented at all levels of the organization and in using the results of global evaluations for strategic planning. These initiatives hold out the promise of strengthened results reporting which could negate the need for development effectiveness reviews of this type in the future. That could depend on how well the evaluation function is able to be incorporated into the system as a means of verifying UNICEF’s contribution and testing the validity of theories of change.

UNICEF has undertaken a number of initiatives in the period following publication of the evaluations reviewed in order to address some of the issues reported here. Examples of the main initiatives include:

- The development and ongoing implementation of new systems for monitoring and reporting results;
- The continued development by the Evaluation Office of the Global Evaluation Report Oversight System (GEROS), coupled with efforts to improve evaluation coverage;
- Changes to operating procedures to reduce procedural bottlenecks;
- The development of the Strategic Priority for Gender Equality Action Plan; and,
- Efforts to use evaluation and research findings to inform and strengthen strategic planning.

While these have not been assessed, documents and interviews at UNICEF suggest that they represent a significant effort to respond to the issues identified. Their effectiveness will doubtless be the subject of future evaluations.
### Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCC</td>
<td>Core Commitments to Children in Humanitarian Emergencies</td>
</tr>
<tr>
<td>CEDAW</td>
<td>Convention on the Elimination of All Forms of Discrimination Against Women</td>
</tr>
<tr>
<td>CIDA</td>
<td>Canadian International Development Agency</td>
</tr>
<tr>
<td>CO</td>
<td>Country Office</td>
</tr>
<tr>
<td>CRC</td>
<td>Convention on the Rights of the Child</td>
</tr>
<tr>
<td>CSO</td>
<td>Civil society organization</td>
</tr>
<tr>
<td>DAC</td>
<td>Development Assistance Committee</td>
</tr>
<tr>
<td>DAC-EVALNET</td>
<td>DAC Network on Development Evaluation</td>
</tr>
<tr>
<td>DFID</td>
<td>Department for International Development (United Kingdom)</td>
</tr>
<tr>
<td>GEROS</td>
<td>Global Evaluation Report Oversight System</td>
</tr>
<tr>
<td>LDC</td>
<td>Least-developed country</td>
</tr>
<tr>
<td>M&amp;E</td>
<td>Monitoring and evaluation</td>
</tr>
<tr>
<td>MDG</td>
<td>Millennium Development Goals</td>
</tr>
<tr>
<td>MIC</td>
<td>Middle-income country</td>
</tr>
<tr>
<td>MOPAN</td>
<td>Multilateral Organizations Performance and Assessment Network</td>
</tr>
<tr>
<td>MoRES</td>
<td>Monitoring Results for Equity Systems</td>
</tr>
<tr>
<td>MTSP</td>
<td>Medium-term Strategic Plan</td>
</tr>
<tr>
<td>NGO</td>
<td>Non-governmental organization</td>
</tr>
<tr>
<td>RBM</td>
<td>Results-based management</td>
</tr>
<tr>
<td>UN</td>
<td>United Nations</td>
</tr>
<tr>
<td>UNAIDS</td>
<td>Joint United Nations Programme on HIV/AIDS</td>
</tr>
<tr>
<td>UNDAF</td>
<td>United Nations Development Assistance Framework</td>
</tr>
<tr>
<td>UNICEF</td>
<td>United Nations Children’s Fund</td>
</tr>
<tr>
<td>UNGEI</td>
<td>United Nations Girls’ Education Initiative</td>
</tr>
<tr>
<td>VISION</td>
<td>Virtual Integrated System of Information</td>
</tr>
<tr>
<td>WASH</td>
<td>Water, sanitation and hygiene</td>
</tr>
</tbody>
</table>
1.0 Introduction

1.1 Background

This report presents the results of a review of the effectiveness of development and humanitarian programming supported by the United Nations Children’s Fund (UNICEF). The common approach and methodology for reviews of this type were developed under the guidance of the Organisation for Economic Co-operation and Development’s Development Assistance Committee (DAC) Network on Development Evaluation (DAC-EVALNET). The review relies on the content of published evaluation reports produced by UNICEF supplemented with a review of UNICEF corporate documents and consultation with staff at UNICEF headquarters in New York.

The method uses a common set of assessment criteria derived from the DAC’s evaluation criteria (Annex 1). It was pilot tested during 2010 using evaluation material from the Asian Development Bank and the World Health Organization. The overall approach and methodology were endorsed by the members of the DAC-EVALNET as an acceptable approach for assessing the development effectiveness of multilateral organizations in June 2011. The first full reviews using the approved methodology were conducted in 2011/2012. The Canadian International Development Agency (CIDA) led the review of the United Nations Development Fund and the World Food Programme, with support and participation from the Netherlands Ministry of Foreign Affairs.

In 2012/13, another two development effectiveness reviews were carried out. CIDA took the lead role in an assessment of the African Development Bank while the Netherlands Ministry of Foreign Affairs was responsible for the review of UNICEF. The review was carried out by an independent team of evaluators from Goss Gilroy Inc. in Ottawa.

The review of UNICEF took place from October 2012 to May 2013. It included three rounds of consultations with UNICEF, including two days of meetings at project startup, a presentation and review of preliminary findings held in March 2013 and the presentation and discussion of the final report on May 6, 2013. The final report was also presented and discussed at a meeting of interested UNICEF donor agencies, hosted by the Netherlands Permanent Mission to the United Nations in New York City on May 7, 2013.

From its beginnings, the process of developing and implementing the reviews of development effectiveness has been coordinated with the work of the Multilateral Organization Performance Assessment Network (MOPAN). By focusing on development effectiveness and carefully selecting assessment criteria, the reviews seek to avoid duplication or overlap with the MOPAN process.
The intent behind the initiative is that EVALNET members engage in development reviews of multilateral organizations as distinct assessments in parallel to the MOPAN reviews. The planning of the reviews should take place in consultation with the evaluation department of the respective multilateral organization. The results of the development reviews will then be of use for the multilateral organizations and their stakeholders.

1.2 Purpose

The purpose of the review is to provide an independent, evidence-based assessment of the humanitarian and development effectiveness of UNICEF operations (hereafter referred to as "programs") for use by all stakeholders.

The current approach to assessing the development effectiveness of multilateral organizations was developed in order to address a gap in the information available to bilateral development agencies. While MOPAN provides regular reports on the organizational effectiveness of multilateral organizations, only in 2012 has it begun to examine development effectiveness and has not yet fully addressed the information gap this review is meant to fill. Other options such as large-scale, joint donor-funded evaluations of a given multilateral organization are much more time-consuming, more costly and result in a significant management burden to the organization being evaluated before, during and after such evaluations.

The current approach is intended to work in a coordinated way with initiatives such as the DAC-EVALNET/United Nations Evaluation Group Peer Reviews of United Nations organization evaluation functions. It also recognizes that multilateral organizations continue to make improvements and strengthen their reporting on development effectiveness. The ultimate aim of the approach is to be replaced by regular, evidence-based, field-tested reporting on development effectiveness provided by multilateral organizations themselves.

1.3 Structure of the Report

The report is structured as follows:

- Section 1.0 provides an introduction to the review and a general description of UNICEF as an organization;
- Section 2.0 presents a brief description of the approach and methodology used to carry out the review;
- Section 3.0 details the findings of the review in relation to six main criteria and
nineteen sub-criteria of effectiveness in development programming and humanitarian action. Each section (one for each of the six main criteria) begins with a discussion of the level of coverage found for each sub-criterion. This is necessary to ensure transparency and that the reader understands the context for the findings. The short sub-sections on coverage are then followed by a detailed discussion of the results found in the evaluations, as they apply to the sub-criteria; and,

- Section 4.0 provides the conclusions of the review. As this is an external review of effectiveness, rather than an evaluation of UNICEF programs, the report does not include recommendations for UNICEF.

1.4 UNICEF: A Global Organization Focused on Equity for Children

1.4.1 UNICEF’s Strategic Direction and Focus on Equity

UNICEF’s mission statement reiterates its mandate from the United Nations to “advocate for the protection of children’s rights, to help meet their basic needs and to expand their opportunities to reach their full potential.” Since 2006 it has been guided in carrying out this mission by a Medium-Term Strategic Plan (MTSP) which originally covered the period from 2006 to 2009, but has twice been extended and now is referred to as the MTSP for 2006-2013.

The MTSP sets out 15 guiding principles, five program focus areas and two foundation strategies. The guiding principles link UNICEF’s programming to normative documents such as the Convention on the Rights of the Child (CRC), the Convention on the Elimination of All Forms of Discrimination Against Women (CEDAW) and the Millennium Declaration and Millennium Development Goals (MDG). They also commit the organization, inter alia, to a strategy of capacity development, working firmly within the United Nations system, and adapting program

---

Box 1: Five UNICEF Program Focus Areas

1. Young child survival and development
2. Basic education and gender equality
3. HIV/AIDS and children
4. Child protection from violence, exploitation and abuse
5. Policy advocacy and partnerships for children’s rights

Medium-Term Strategic Plan, 2006 – 2013

---


strategies to the needs of program countries.

The five program focus areas are a central organizing principle of the MTSP (Box 1). They are intended to “have a decisive and sustained impact on realizing children’s rights and achieving the commitments of the Millennium Declaration and Goals.”

The MTSP provides a description of the types of activities UNICEF will undertake and the programs it will support in each of its five focus areas (Figure 1).

**Figure 1: UNICEF Activities and Programs in Five Focus Areas of the MTSP**

<table>
<thead>
<tr>
<th>UNICEF MTSP Focus Areas</th>
<th>UNICEF Program Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Young child survival and development</td>
<td>Support in regular, emergency and transitional situations for essential health, nutrition, water and sanitation programmes, and for young child and maternal care at the family, community, service-provider and policy levels.</td>
</tr>
<tr>
<td>2. Basic education and gender equality</td>
<td>Focus on improved developmental readiness for school; access, retention and completion, especially for girls; improved education quality; education in emergency situations and continued leadership of the United Nations Girls’ Education Initiative (UNGEI).</td>
</tr>
<tr>
<td>3. HIV/AIDS and children</td>
<td>Emphasis on increased care and services for children orphaned and made vulnerable by HIV/AIDS, on promoting expanded access to treatment for children and women and on preventing infections among children and adolescents; continued strong participation in the Joint United Nations Programme on HIV/AIDS (UNAIDS).</td>
</tr>
<tr>
<td>5. Policy advocacy and partnerships for children’s rights</td>
<td>Putting children at the centre of policy, legislative and budgetary provisions by: generating high-quality, gender-disaggregated data and analysis; using these for advocacy in the best interests of children; supporting national emergency preparedness capacities; leveraging resources through partnerships for investing in children; and fostering children’s and young people’s participation as partners in development.</td>
</tr>
</tbody>
</table>

Source: Medium-Term Strategic Plan 2006 - 2009

In addition to the five focus areas, the MTSP re-iterates UNICEF’s intention to build its capacities to respond to emergencies in a timely and effective manner. It notes that responding to humanitarian situations remains an essential element of the work of UNICEF.

---

3 MTSP, 2006-2013, p.16.
Finally, the MTSP includes a major commitment to the pursuit of equity by focusing on gender equality and the rights of the most vulnerable. It identifies a human-rights based approach to programming and promoting gender equality as foundation strategies for UNICEF.

Applying a human rights-based approach and promoting gender equality, as “foundation strategies” for UNICEF work will improve and help to sustain the results of development efforts to reduce poverty and reach the Millennium Development Goals by directing attention, long-term commitment, resources and assistance from all sources to the poorest, most vulnerable, excluded, discriminated and marginalized groups.\(^4\)

The principles, focus areas and foundation strategies outlined in UNICEF’s MTSP have been ratified by two Mid-Term Review Reports (2008 and 2010).\(^5\)

The most recent annual report of Executive Director of UNICEF placed special emphasis on the organizations “refocus on equity”:

The refocus on equity holds abundant promise for children, especially through faster and more economical achievement of the MDGs. UNICEF emphasized the implementation of the refocus during 2011 at the local and national levels in partnership with governments, civil society organizations and United Nations partners. Measuring results for the most disadvantaged populations, particularly at the local level, has proven critical to accelerate and sustain progress in reducing disparities.\(^6\)

Thus UNICEF’s strategy for 2006-13 has placed considerable emphasis on addressing inequities faced by vulnerable children through programming in its five program focus areas, responding effectively to humanitarian situations and pursuing foundation strategies of a human rights-based approach to programming and promoting gender equality.

1.4.2 UNICEF Operations and Program Expenditures

UNICEF is active in over 190 countries and territories through its programs and the work of national committees. In 2011 it had programs of cooperation in 151 countries, areas and territories: 45 in sub-Saharan Africa; 35 in Latin America and the Caribbean; 35 in Asia; 16 in the Middle East and North Africa; and 20 in Central and Eastern Europe and

---

\(^4\) MTSP, 2006-13, p.31.


UNICEF’s spending on program assistance in 2011 totaled $3,472 million. Over half that amount went to the Young Child Survival and Development focus area. Similarly 57\% of total program assistance in 2011 was directed to sub-Saharan Africa which has the majority of the world’s Least Developed Countries (LDCs).\footnote{Ibid., p.9.}

Figure 2 shows the portion of UNICEF direct program expenditures in each focus area and in humanitarian action from 2009 to 2011.

**Figure 2: UNICEF Expenditures, by Focus Area and Humanitarian Action, 2009-2011**


Expenditures in the five focus areas are development expenditures funded with a) Regular Resources and b) Regular Resources – Other, using UNICEF’s system for accounting for resources. Expenditures on humanitarian action are those financed by c) Other Resources – Emergency. The distribution of expenditures by type of funding from 2009 to 2011 is shown in Figure 3.

\footnotesize
\begin{itemize}
  \item 1. Young child survival and development
  \item 2. Basic education and gender equality
  \item 3. HIV/AIDs and children
  \item 4. Child protection
  \item 5. Policy advocacy and partnership
  \item Other
  \item Humanitarian Emergency
\end{itemize}
1.5 Program Evaluation at UNICEF

One of the conditions for undertaking a development effectiveness review is the availability of enough evaluation reports of reasonable quality to provide an illustrative sample covering an agency’s activities and programs (Annex 2). In order to satisfy that condition the review team examined program evaluation policies and practices and the quality of evaluation reports produced and published by UNICEF. This short survey is by no means exhaustive and should not be read as an overall assessment of the function. It is only intended to establish the feasibility of conducting a development effectiveness review of UNICEF based on its own published evaluations.

1.5.1 Evaluation Policies and Practices at UNICEF

UNICEF’s current evaluation policy⁹ was prepared, at least in part, in response to the report of a peer review panel of international evaluation experts produced under the auspices of the DAC-EVALNET and the United Nations Evaluation Group in 2006. That report was just the second in the ongoing series of peer reviews of evaluation functions in the multilateral system. In its judgment statement, the report recognized the strength of UNICEF’s Evaluation Office and the ongoing challenge of ensuring evaluation quality and adequate coverage in a decentralized system of program evaluation:

Evaluation at UNICEF is highly useful for learning and decision-making.

---

purposes and, to a lesser extent, for accountability in achieving results.

UNICEF’s central Evaluation Office is considered to be strong, independent and credible. Its leadership by respected professional evaluators is a major strength. The EO has played an important leadership role in UN harmonization through the UN Evaluation Group.

The Peer Review Panel considers that a decentralized system of evaluation is well-suited to the operational nature of UNICEF. However, there are critical gaps in quality and resources at the regional and country levels that weaken the usefulness of the evaluation function as a management tool.\(^{10}\)

The 2008 evaluation policy for UNICEF retained the decentralized system wherein responsibility for different types of evaluations rested at different levels in the organization. Figure 4 presents an overview of evaluation roles and responsibilities by organizational location at UNICEF as described in the policy.

**Figure 4: Organizational Responsibilities for Program Evaluation at UNICEF**

<table>
<thead>
<tr>
<th>Organizational Location</th>
<th>Evaluation Roles and Responsibilities</th>
<th>Types of Evaluations Commissioned</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country Offices/Country Representatives</td>
<td>• Assign resources to evaluation&lt;br&gt;• Communicate with partners&lt;br&gt;• Prepare Integrated Monitoring and Evaluation Plan for the country office&lt;br&gt;• Provide quality assurance for meeting standards established by the Evaluation Office&lt;br&gt;• Ensure evaluation findings inform decision making process&lt;br&gt;• Follow up and report on the status of evaluation recommendations</td>
<td>• Local or project evaluations&lt;br&gt;• Program evaluations of elements of the country office program</td>
</tr>
<tr>
<td>Regional Offices/Regional Directors</td>
<td>Focuses on oversight and strengthening the capacity of country office evaluation functions by:&lt;br&gt;• Coordinating capacity building with the Evaluation Office&lt;br&gt;• Preparing regional evaluation plans&lt;br&gt;• Provide quality assurance and technical assistance to evaluations of country programs</td>
<td>• Country program evaluations&lt;br&gt;• Multi-country thematic evaluations&lt;br&gt;• Regional contribution to global evaluations&lt;br&gt;• Real-time evaluations of emergency operations</td>
</tr>
<tr>
<td>Directors at Headquarters</td>
<td>• Prioritize evaluations for global policies and initiatives in their program areas.</td>
<td>• Global policies&lt;br&gt;• Global program initiatives</td>
</tr>
</tbody>
</table>

### Organizational Location

<table>
<thead>
<tr>
<th>Evaluation Roles and Responsibilities</th>
<th>Types of Evaluations Commissioned</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Ensure funding for programs funded by other resources.</td>
<td>• Independent global evaluations</td>
</tr>
<tr>
<td>Evaluation Office</td>
<td>• Conducts evaluations of global policies and program initiatives as requested by Directors.</td>
</tr>
<tr>
<td>• Coordinates evaluation function at UNICEF</td>
<td></td>
</tr>
<tr>
<td>• Collaborates with UNICEF partners in multi-party evaluations</td>
<td></td>
</tr>
<tr>
<td>• Promotes capacity development in evaluation in developing countries</td>
<td></td>
</tr>
<tr>
<td>• Provides leadership in development of approaches and methodologies for policy, strategic, thematic, programme project and institutional evaluations</td>
<td></td>
</tr>
<tr>
<td>• Maintains the institutional database of evaluations and promotes its use</td>
<td></td>
</tr>
<tr>
<td>• Conducts periodic meta-evaluations of the quality and use of evaluations at UNICEF</td>
<td></td>
</tr>
<tr>
<td>Evaluation Committee</td>
<td></td>
</tr>
<tr>
<td>• Reviews UNICEF evaluations with relevance at global governance level</td>
<td></td>
</tr>
<tr>
<td>• Examines annual follow up reports on the implementation of recommendations</td>
<td></td>
</tr>
<tr>
<td>• Reviews the work program of the Evaluation Office</td>
<td></td>
</tr>
</tbody>
</table>


UNICEF reports annually to the Executive Board on the performance of the evaluation function. This report encompasses both the number and types of evaluations produced and progress in monitoring and improving the quality of the evaluations produced.

For the past three years, the quality of UNICEF evaluations produced by country and regional offices has been assessed independently by a contracted external agency applying a consistent methodology. The results of the Global Evaluation Report Oversight System (GEROS) are published annually. The results of the GEROS annual review are incorporated into the Annual report on the evaluation function. Figure 5 summarizes the evaluation results indicators reported in the 2012 report.

**Figure 5: Performance Indicators for the Evaluation Function at UNICEF 2011**

<table>
<thead>
<tr>
<th>Evaluation Performance Indicator</th>
<th>Results for 2011</th>
</tr>
</thead>
</table>
| 1. Number of evaluations managed and submitted to the Global Database | • 140 evaluations  
• 700 studies  
• Avg. = 1 evaluation and 7 studies per office |
| 2. Topical distribution | • 25% Child survival and development  
• 15% Education and gender |
### Evaluation Performance Indicator | Results for 2011
---|---
| | 10% HIV/AIDS
| | 16% Child protection
| | 1% Policy, advocacy and partnership
| | 33% Cross cutting themes and organizational performance
3. Type of evaluations conducted | 33% Impact
| | 26% Outcomes
| | 33% Outputs
4. Quality of UNICEF evaluations | GEROS Rating for evaluations assessed in 2011, conducted in 2010
| | 40% satisfactory
| | 30% nearly satisfactory
| | 30% poor quality
5. Use of evaluation, including management responses | 65% of evaluations upload management response to tracking system
6. Corporate-level evaluations | 13 corporate level evaluations (from a target of 15) completed or in progress by April 2012

Source: *Annual Report on the evaluation function and major evaluations at UNICEF*, p.3 - 5.

The annual GEROS report represents an important effort by the Evaluation Office to monitor and influence the quality of evaluations produced by regional and country offices. This has been further strengthened by the development of a dashboard for summarizing information on the decentralized evaluation function (regional and country office level) which is part of UNICEF’s Virtual Integrated System of Information (VISION) launched in January 2012.

The decentralized evaluation function dashboard is uploaded quarterly and presents (for each UNICEF regional office) information on: evaluation report submission rates; the quality of evaluation reports as assessed by GEROS in the previous year; management responses uploaded to the Global Tracking System; and, the implementation rates for management responses.

What is interesting is the absence of a metric used to measure coverage of country office programs in the VISION dashboard for decentralized evaluation. The review team found (Section 2.0) considerable variation in the number of evaluations carried out over three years (2009 to 2011) by the country offices. Nonetheless, it is clear that UNICEF’s Evaluation Office is engaged in ongoing efforts to upgrade the quality of decentralized evaluations.

Finally, it is worth pointing out that recent global evaluations managed by the Evaluation Office have been used by Programme Division at UNICEF to inform the planning.
1.5.2 Quality of Evaluation Reporting at UNICEF

Data on evaluation quality at UNICEF is available for the review from two sources: UNICEF’s independent GEROS reports and the results of the review team’s quality assessment of evaluations in the sample.

The latest report of the GEROS was produced in 2012 and covers 87 evaluation reports produced by UNICEF in 2011. It reports a three-year improving trend with 42% of reports meeting UNICEF evaluation standards in 2011 versus 40% in 2010 and 36% in 2009.\(^\text{12}\)

The review took the GEROS ratings into account when developing the evaluation sample as noted in Annex 3. There are four evaluation report quality ratings used in the GEROS system: Not confident to act; Almost confident to act; Confident to act; and, Very confident to act. In developing a purposive sample of the evaluation reports produced from 2009 to 2011, the review team included only those evaluations rated almost confident to act or better. The review team assessed a total of 70 evaluations for quality and rejected only four for quality reasons. Four others were not included in the review for other reasons (Section 2.1 and Annex 3).

1.6 Monitoring and Reporting Results

1.6.1 Elements of UNICEF’s Results Monitoring and Reporting Systems

In recent years, UNICEF has invested considerable effort in modifying and strengthening its systems and processes for monitoring and reporting on the results of programs, and for allocating resources to specific results. Currently, some of the most notable elements of the system include the Monitoring Results for Equity Systems (MoRES), the Virtual Integrated System of Information (VISION) and the Executive Director’s annual reports on progress and achievement against the MTSP.

Monitoring Results for Equity Systems

The MoRES system is one of the core elements in UNICEF’s “refocus on equity” which began in 2010 as an initiative of the Executive Director. MoRES is UNICEF’s conceptual


framework for planning, programming, implementation and management of results.\textsuperscript{13} It is intended to address the need for intermediate process and outcome measures between the routine monitoring of program and project inputs and outputs on one hand and the monitoring of higher level outcomes; which is done using internationally agreed indicators every three to five years.

The conceptual model for MoRES includes four levels of country office programmatic actions, each linked to different planning, monitoring, and evaluation activities. Level three monitoring, which is the primary focus of MoRES, is focused on the question: “Are we going in the right direction? Are bottlenecks and barriers to equity changing?”\textsuperscript{14}

In order to undertake level three monitoring, country offices must first identify country specific indicators relating to ten determinants of bottlenecks hindering equity for children. These determinants are gathered under four major headings: the quality, demand for, and supply of services; and the enabling environment.

In 2012, the system of indicators for results in humanitarian action at UNICEF, the Humanitarian Performance Monitoring System, was incorporated into MoRES to allow for a single, integrated system of performance monitoring.

It is worth noting that the conceptual model of MoRES includes the evaluation function as the means of verifying final outcomes and impacts of programming at country level.

The global two year management plan for MoRES was adopted by UNICEF in January 2012, and the level three monitoring module was implemented by 27 country offices in the March-June period of the same year\textsuperscript{15}. Mainstreaming of MoRES into the annual review process took place in the last half of 2012. As an important element of the refocus on equity, MoRES is to be the subject of a formative evaluation managed by the Evaluation Office in 2013.

\textbf{Virtual Integrated System of Information}

The VISION system, fully launched in January 2012, allows UNICEF offices at all levels to develop and input data on different types and levels of results. Perhaps most importantly it links together budgeting data and key results so that country offices can now input their budget information under key UNICEF results areas rather than program expenditure activities.

\textsuperscript{13} MoRES: Monitoring Results for Equity Systems: Access and Quality for Early Learning, UNICEF, 2012.


A key component of VISION is its Results Assessment Module which can be used to extract information on program planned results and compare those to reported achievements. During 2011, MoRES level three monitoring elements were integrated into VISION.

**Annual Report of the Executive Director of UNICEF: progress and achievement against the medium term plan**

The Executive Director of UNICEF reports annually on the progress made and achievements secured against the key results identified in the MTSP. This report combines information on global trends in indicators relating to children-specific MDGs (and other global, regional and country-specific indicators of child welfare and disparity) with examples of UNICEF’s contribution to results in each of the five focus areas.

The annual report also addresses developments in the implementation of programme principles and strategies and on internal measures of programme performance. In 2012, the report highlighted developments in evaluation, country office efficiency, shared services with other UN agencies, and new initiatives in recruiting and human resource management.

The annual report is supported by a large data companion report which tracks a significant number of key performance indicators across the five focus areas and humanitarian action.

**1.6.2 Strengths and Weaknesses of the Results Monitoring and Reporting System**

The main strength of UNICEF’s results monitoring and reporting system as it currently stands is its multi-layered nature and its ability to link standardized categories of results to resource allocations and budgets through the VISION system. With MoRES level three monitoring integrated into the VISION system there should be strong linkages from development and humanitarian program inputs to the intermediate level of results.

What is lacking currently from the annual report of the Executive Director and its data companion is a clear explanation of how UNICEF’s activities and its support to programs contribute to the higher level outcomes and impacts being reported. The contribution is asserted and examples are given, but it is not clear how the link between UNICEF’s contribution and the outcomes is established. At the same time it is important to recognize that MoRES by focusing on determinants of inequality for children and addressing indicators of intermediate results for UNICEF-supported programs goes some way in addressing the issue of UNICIF’s contribution.

The need for strengthened monitoring is recognized in Programme Division’s recent
work on the challenges of the upcoming MTSP.\textsuperscript{16}

Monitoring and demonstrating tangible results is also central for UNICEF’s results and performance-based management and reporting to donors, particularly in the light of increasing competition for humanitarian and development funding. The expansion of the assessment, monitoring, analysis and evaluation functions across country offices and Government partners will require additional investments in staff, training and technical assistance. The next MTSP should emphasize monitoring as a key program function and highlight the investments needed to improve monitoring functions on the ground.

One development which may potentially strengthen the link between higher level outcomes and UNICEF’s contribution is the explicit emphasis in MoRES of the role which evaluation can play in verifying outcomes and impacts of country level programming. This will depend, of course, on the development of credible theories of change for country programs and their elements. To some extent, the MoRES model addresses this question through its identification of determinants of equity. However, this is not yet made explicit in reporting on the progress against the MTSP.

In summary, UNICEF is investing considerable time and effort into the development of quite new systems for results monitoring which are being implemented at all levels of the organization. They hold out the promise of strengthened results reporting which could negate the need for development effectiveness reviews of this type in the future. That could depend on how well the evaluation function is able to be incorporated into the system as a means of verifying UNICEF’s contribution and testing the validity of theories of change.

2.0 Methodology

This section describes the methodology used for the review, including the identification of the population and the sampling process, the review criteria, the review and analysis processes and other data collection used. It concludes with a discussion of the limitations of the review.

2.1 Evaluation Population and Sample

A population of UNICEF evaluations was identified from two sources:

- Office of Evaluation website (www.unicef.org/evaldatabase); and,
- Three oversight reports from the assessment of evaluations in UNICEF’s Global Evaluation Reports Oversight System (GEROS), covering evaluations conducted from 2009 to the end of 2011.

These sources identified a population of 341 evaluations. It was decided, in consultation with UNICEF’s Office of Evaluation, that the review should focus on evaluations conducted since 2008. UNICEF had adopted a new evaluation policy in that year and it was expected that evaluations after 2008 would reflect this new policy. It was also determined that the sample would not include those evaluations deemed, in the GEROS quality rating, to be of poor quality and “not confident to act.” Evaluations that focused on global programming were not included in the population of evaluations for the quantitative review, but included in a qualitative review by senior team members. In addition, some evaluations in the population were duplicates, were not deemed to be evaluations or did not focus on UNICEF programming and were dropped from the population. This resulted in a population of 197 from which to draw the sample.

A sample of 70 evaluations was drawn from the population, with the intention of having 60 evaluations that would pass the quality review and be available for rating. Initially the sample was random, stratified by MTSP theme area. Then the sample was adjusted manually to ensure adequate coverage on two other dimensions: region and type of country (least-developed or middle-income countries). The resulting sample was no longer random, but rather a purposive sample that is illustrative of UNICEF’s programming across a number of dimensions. Eight evaluations were eliminated from the initial sample – four because they did not pass the quality review, two because they did not cover UNICEF programming or address program effectiveness and two because they covered activities that were already covered in other evaluations. This left a sample of 62 evaluations that were rated by the team (Annex 2).
The key characteristics of the population and the sample include:

- Evaluations in the population and sample were evenly spread across the three years 2009 – 2011;

- Evaluations in the population and sample were split among the regions in which UNICEF is programming and between middle- and least-developed countries. However, in the sample there were slightly more evaluations of programming in least-developing countries in West and Central Africa than were represented in the population;

- Evaluations in the population and sample covered all priority areas of UNICEF programming, including all themes in the MTSP and humanitarian action. However, not all types of UNICEF programming are covered equally by evaluations in the population. Relatively few evaluations in the population covered humanitarian action; whereas this programming accounted for just over one-quarter of UNICEF expenditures in 2009 – 2011. This is also reflected in the fact that, of the countries receiving the largest amounts of humanitarian funding in 2009 – 2011, very little of their humanitarian action was subject to evaluation. In addition, five of the humanitarian evaluations included in the sample were for programming in response to the same humanitarian crisis – the 2004 Indian Ocean tsunami;

- The evaluation population and sample also under-represented evaluations of humanitarian actions managed by UNICEF at the Country Office level (Level One), which are quite numerous. All but one of the evaluations of humanitarian action included in the sample covered UNICEF Level Three emergencies. Thus the findings on humanitarian action are all concerned with emergencies managed by UNICEF at the global or regional level.

- Only a small percentage of UNICEF’s evaluations cover programming in the countries receiving the largest amounts of development or humanitarian funding in 2007 - 2011. It is recognized that the evaluations cover funding periods that span from 2000 to 2011 and the profile of spending in the years prior to 2007 may have been different. However, the under-representation of the largest funded countries in the population does suggest that UNICEF’s is not achieving strong coverage of its largest amounts of development or humanitarian funding. The sample attempted to compensate for this somewhat, by including a larger number

---

17 UNICEF responded to 232 emergencies in 2009; 290 in 2010 and 292 in 2011. Most of these were Level One emergencies. (UNICEF communication).
of evaluations from these countries than would be expected for the distribution of
the population.\textsuperscript{18}

See Annex 3 for details of the comparison of the evaluation population and sample.

It was not possible to identify the value of UNICEF programming that was covered by
the evaluation sample. While the review process required the raters to identify the overall
costs of the programs, this information was found in only 34 evaluations and the nature of
the information presented was inconsistent across evaluations.\textsuperscript{19}

Apart from the fact that the evaluations do not cover the countries receiving the largest
amounts of funding, or the full range of humanitarian action, the evaluations covered in
the sample are, in all other aspects, illustrative of UNICEF’s global programming.

\subsection*{2.2 Criteria}

The methodology does not rely on a particular definition of development and
humanitarian effectiveness. As agreed with the Management Group and the Task Team
that were created by the DAC-EVALNET to develop the methodology of these reviews,
the methodology focuses on some essential characteristics of effective multilateral
organization programming, derived from the DAC evaluation criteria:

\begin{itemize}
\item Programming activities and outputs would be relevant to the needs of the target
group and its members;
\item The programming would contribute to the achievement of development objectives
and expected development results at the national and local level in developing
countries (including positive impacts for target group members);
\item The benefits experienced by target group members and the development (and
humanitarian) results achieved would be sustainable in the future;
\item The programming would be delivered in a cost efficient manner;
\item The programming would be inclusive in that it would support gender equality and
would be environmentally sustainable (thereby not compromising the
development prospects in the future); and,
\item The programming would enable effective development by allowing participating
\end{itemize}

\textsuperscript{18} This explains the proportionately larger number of evaluations from least-developing countries in West
and Central Africa.

\textsuperscript{19} Sometimes reflecting the overall costs, at some only UNICEF’s costs; at times reflecting the budget, at
times the actual expenditures.
and supporting organizations to learn from experience and use of performance management and accountability tools, such as evaluation and monitoring to improve effectiveness over time.

The review methodology, therefore, involves a systematic and structured review of the findings of UNICEF evaluations, as they relate to six main criteria (Box 2) and 19 sub-criteria that are considered to be essential elements of effective development and humanitarian programming (Annex 1).

**Box 2: Assessment Criteria**

1. Relevance of Interventions
2. The Achievement of Humanitarian and Development Objectives and Expected Results
3. Cross Cutting Themes (Environmental Sustainability and Gender Equality)
4. Sustainability of Results/Benefits
5. Efficiency
6. Using Evaluation and Monitoring to Improve Humanitarian and Development Effectiveness

### 2.3 Review Process and Data Analysis

Each evaluation was reviewed by one member of a small review team that included three reviewers and two senior members (including the team leader). Each team member reviewed a set of evaluations. The first task of the reviewer was to assess the quality of the evaluation report to ensure that it was of sufficiently high quality to provide reliable information on development and humanitarian effectiveness. This was done using a quality scoring grid (Annex 3). If the evaluation met the minimum score required, the reviewers continued to provide a rating on each sub-criterion, based on information in the evaluations and standard review grid (Annex 4). They also provided evidence from the evaluations to substantiate the ratings.

All efforts were made to ensure consistency in the ratings by team members. The reviewers were trained by the senior team members; two workshops were held at which all team members reviewed and compared the ratings for the same three evaluations, a mid-term workshop was held to address any issues faced by the reviewers, and, following completion of the reviews and the documentation of the qualitative evidence to support the ratings, the team leader reviewed all ratings to ensure there was sufficient evidence provided and it was consistent with the rating. Senior team members then reviewed the

---

20 All team members have graduate degrees in fields related to public policy, evaluation and/or international development. The two senior members have more than twenty-five years experience in evaluation, including international development evaluation. One team member has over twenty years of experience in development programming, with specific expertise in gender-based issues. The other team members have between one and four years of experience with international development projects.
qualitative evidence for each sub-criterion to identify factors contributing to, or detracting from, the achievement of the sub-criteria.

2.4 Other Data Collection

After the rating of the evaluations in the sample, further information about UNICEF’s effectiveness was also gleaned from a qualitative review of the global evaluations conducted in the 2009–2011 period. These were not included in the sample for rating in order to avoid double-counting results. Some global evaluations included evidence drawn from other evaluations and, as such, there was a risk that evaluative information on the same program would be counted twice. The separate review of the global evaluations also gave them a higher profile in the analysis of the findings for the six criteria.

The review of evaluation reports was also supplemented by a review of UNICEF corporate documents and interviews with UNICEF staff. These were done to contextualize the results of the review. (A list of the global evaluations and documents consulted is provided in Annex 5.)

2.5 Limitations

As with any meta-synthesis, there are methodological challenges that limit the findings. For this review, the limitations include sampling bias, the challenge of assessing overall programming effectiveness when there are important variations in programming (for example, evaluations covering multiple programming components or only a specific theme or project as part of a program area) and the retrospective nature of a meta-synthesis.

Sampling Bias

The sample selected for this review is not a random sample. Even as a purposive sample, it is still illustrative of UNICEF-supported programming. However, caution must be exercised in generalizing from the findings of this review to all UNICEF programming. In addition, although the UNICEF’s evaluations cover a range of countries and all focus areas, as outlined in the MTSP, there are fewer evaluations of programming in countries receiving the largest amounts of UNICEF funding and its humanitarian action than would be expected, given its proportion of overall funding at UNICEF. However, the review did attempt to compensate for this by over-sampling from countries with the largest amounts of funding and those with humanitarian crises. As already noted, neither the population nor the sample included evaluations of small (Level One) humanitarian actions at UNICEF.
There was generally adequate coverage of the criteria. Seventeen of the 19 sub-criteria used to assess development and humanitarian effectiveness are sufficiently well covered in the evaluations reviewed to be included in this report. Two received a weak coverage rating and the results are not reflected in the report.

**Variations in Programming**

The review was not able to report systematically on the effectiveness of UNICEF’s programming by focus area or by type of country. There were not sufficient evaluations in each focus area or type of country included in the meta-synthesis to allow for them to be analyzed separately. In addition, some evaluations cover multiple types of programming in the same evaluation. This means that it is not possible to draw conclusions by type of programming or type of country. Although not analyzed separately, where qualitative observations can be made about the focus area and/or the country type, these are reflected in the report. However, they should be treated with caution, as they can only be illustrative. The quantitative findings, where appropriate, for these analyses are reported in Annex 6.

**Retrospective Nature of Meta-synthesis**

Evaluations are, by definition, retrospective and a meta-synthesis is even more retrospective, as it is based on a body of evaluations that address policies and programming implemented over a much earlier period of time. UNICEF’s evaluations published in 2009 – 2011, covered years beginning in 2000 and running to 2011.\(^{21}\) UNICEF’s policies, strategies and approaches to programming have changed over these years, but the changes will not be reflected in all the evaluations. For this reason, the findings may be somewhat dated. To the extent possible, the review addressed this through observations gleaned from recent interviews with UNICEF staff and a review of UNICEF documents.

---

\(^{21}\) The longest time span was eleven years – 2000 to 2010 – and the most recent a one-year program in 2010.
3.0 Findings on UNICEF’s Development Effectiveness

This section presents the results of the review as they relate to the six main criteria and their associated sub-criteria. For each criterion, the report presents firstly the extent to which the review sub-criterion was addressed in the evaluation reports (coverage). Then the results are presented (findings) – including both the quantitative findings for each sub-criterion and the results of the qualitative analysis of the factors contributing to, or detracting from, the achievement of the sub-criteria. This section includes also evidence from global evaluations that were not included in the quantitative review.

In reporting on the factors, the report makes use of the terms “most”, “many”, “some” and “few” to describe the frequency with which an observation was noted, as a percentage of the number of evaluations for which the sub-criterion was covered (Box 3).

In addition, for the most part, the order in which the factors are presented reflects the frequency with which they were mentioned.

Figure 6 summarizes the findings with respect to coverage (represented by the letter n), the team’s assessment of whether this is strong, moderate or weak, and the ratings assigned by the review team of “satisfactory” or “unsatisfactory” for each of the six major criteria and their associated sub-criteria.

### Box 3: Frequency of Observations

Most = over three-quarters of the evaluations for which the sub-criterion was covered

Many = between half and three-quarters

Some = between one-fifth and half

Few = less than one-fifth

### Figure 6: Coverage and Summary Results for Each Sub-criterion

<table>
<thead>
<tr>
<th>Criteria and Sub-Criteria</th>
<th>n*</th>
<th>Coverage Level**</th>
<th>Satisfactory Ratings (%)***</th>
<th>Unsatisfactory Ratings (%)***</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Relevance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.1 Programs and projects are suited to the needs and/or priorities of the target group.</td>
<td>60</td>
<td>Strong</td>
<td>90%</td>
<td>10%</td>
</tr>
<tr>
<td>1.2 Projects and programs align with national humanitarian and development goals.</td>
<td>61</td>
<td>Strong</td>
<td>98%</td>
<td>2%</td>
</tr>
<tr>
<td>1.3 Effective partnerships with governments, bilateral and multilateral development and humanitarian organizations and Non-governmental organizations for planning, coordination and implementation of support to development and/or emergency preparedness, humanitarian relief and rehabilitation efforts.</td>
<td>56</td>
<td>Strong</td>
<td>88%</td>
<td>13%</td>
</tr>
<tr>
<td><strong>Objectives Achievement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1 Programs and projects achieve their stated humanitarian and development objectives and</td>
<td>61</td>
<td>Strong</td>
<td>77%</td>
<td>23%</td>
</tr>
</tbody>
</table>
## Review of UNICEF’s Development Effectiveness

### Criteria and Sub-Criteria

<table>
<thead>
<tr>
<th>Criteria and Sub-Criteria</th>
<th>n*</th>
<th>Coverage Level**</th>
<th>Satisfactory Ratings (%)***</th>
<th>Unsatisfactory Ratings (%)***</th>
</tr>
</thead>
<tbody>
<tr>
<td>attain expected results.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.2 Programs and projects have resulted in positive benefits for target group members.</td>
<td>62</td>
<td>Strong</td>
<td>85%</td>
<td>15%</td>
</tr>
<tr>
<td>2.3 Programs and projects made differences for a substantial number of beneficiaries and where appropriate contributed to national humanitarian and development goals.</td>
<td>61</td>
<td>Strong</td>
<td>72%</td>
<td>28%</td>
</tr>
<tr>
<td>2.4 Programs contributed to significant changes in national humanitarian and development policies and programs (including for disaster preparedness, emergency response and rehabilitation) (policy impacts) and/or to needed system reforms.</td>
<td>43</td>
<td>Moderate</td>
<td>77%</td>
<td>23%</td>
</tr>
</tbody>
</table>

### Cross-Cutting Themes – Inclusive Humanitarian and Development which is Sustainable

<table>
<thead>
<tr>
<th>Cross-Cutting Themes – Inclusive Humanitarian and Development which is Sustainable</th>
<th>n*</th>
<th>Coverage Level**</th>
<th>Satisfactory Ratings (%)***</th>
<th>Unsatisfactory Ratings (%)***</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Extent to which multilateral organization supported activities effectively address the cross-cutting issue of gender equality.</td>
<td>42</td>
<td>Moderate</td>
<td>48%</td>
<td>52%</td>
</tr>
<tr>
<td>3.2 Extent to which changes are environmentally sustainable.</td>
<td>15</td>
<td>Weak</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

### Sustainability of Results/Benefits

<table>
<thead>
<tr>
<th>Sustainability of Results/Benefits</th>
<th>n*</th>
<th>Coverage Level**</th>
<th>Satisfactory Ratings (%)***</th>
<th>Unsatisfactory Ratings (%)***</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Benefits continuing or likely to continue after project or program completion or there are effective measures to link the humanitarian relief operations, to rehabilitation, reconstructions and, eventually, to longer term humanitarian and development results.</td>
<td>59</td>
<td>Strong</td>
<td>51%</td>
<td>49%</td>
</tr>
<tr>
<td>4.2 Projects and programs are reported as sustainable in terms of institutional and/or community capacity.</td>
<td>59</td>
<td>Strong</td>
<td>47%</td>
<td>53%</td>
</tr>
<tr>
<td>4.3 Programming contributes to strengthening the enabling environment for humanitarian and development.</td>
<td>40</td>
<td>Strong</td>
<td>70%</td>
<td>30%</td>
</tr>
</tbody>
</table>

### Efficiency

<table>
<thead>
<tr>
<th>Efficiency</th>
<th>n*</th>
<th>Coverage Level**</th>
<th>Satisfactory Ratings (%)***</th>
<th>Unsatisfactory Ratings (%)***</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Program activities are evaluated as cost/resource efficient.</td>
<td>48</td>
<td>Strong</td>
<td>63%</td>
<td>38%</td>
</tr>
<tr>
<td>5.2 Implementation and objectives achieved on time (given the context, in the case of humanitarian programming).</td>
<td>39</td>
<td>Moderate</td>
<td>51%</td>
<td>49%</td>
</tr>
<tr>
<td>5.3 Systems and procedures for project/program implementation and follow up are efficient (including systems for engaging staff, procuring project inputs, disbursing payment, logistical arrangements etc.).</td>
<td>35</td>
<td>Moderate</td>
<td>37%</td>
<td>63%</td>
</tr>
</tbody>
</table>

### Using Evaluation and Monitoring to Improve Humanitarian and Development Effectiveness

<table>
<thead>
<tr>
<th>Using Evaluation and Monitoring to Improve Humanitarian and Development Effectiveness</th>
<th>n*</th>
<th>Coverage Level**</th>
<th>Satisfactory Ratings (%)***</th>
<th>Unsatisfactory Ratings (%)***</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1 Systems and process for evaluation are effective.</td>
<td>42</td>
<td>Moderate</td>
<td>64%</td>
<td>36%</td>
</tr>
</tbody>
</table>
## 3.1 Relevance

### 3.1.1 Coverage

There is strong coverage of all sub-criteria with respect to relevance (Figure 7). Almost all evaluations covered the Sub-criterion 1.1 “Programs suited to the needs of the target group” and Sub-criterion 1.2 “Programs aligned with national development goals.” Only slightly fewer evaluations address Sub-criterion 1.3 of “Effective partnerships.”

**Figure 7: Number of Evaluations Addressing Sub-criteria for Relevance**

![Figure 7](image)

### 3.1.2 Key Findings

The findings reflect that UNICEF’s programming is highly relevant to the needs of the target groups (Figure 8). Ninety percent (90%) or more of the evaluations reported suitability to target group needs and alignment with national development goals as...
satisfactory or better. Eighty-eight percent (88%) of the evaluations reported partnerships as being satisfactory or better. Evaluations of programming in middle-income countries (MICs) are somewhat more likely to have satisfactory or better ratings for effective partnerships than those from least-developed countries (LDCs) (Annex 6). Although UNICEF’s programming reflects overall a high level of relevance, the ratings for Sub-criterion 1.1 “Programs suited to the needs of the target group” and Sub-criterion 1.2 “Programs aligned with national development goals” for programming in the area of Child Protection tend to reflect somewhat lower ratings. This likely illustrates the extent to which child protection is not as well recognized, or as well reflected in national policies and priorities, as other sectors of UNICEF programming.

**Figure 8: Findings for Relevance**

![Bar chart showing ratings for relevance](chart.png)

Humanitarian evaluations reflected a somewhat lower rating with respect to being suited to the needs of target groups, perhaps because of the challenges of conducting adequate needs assessments. One evaluation of a major humanitarian response in 2010 noted that there was “virtually no systematic needs assessment that might have informed subsequent cluster coordination and programme design, helped management monitor UNICEF’s progress over time or kept key stakeholders abreast of developments with accurate, reliable information about the situation of women and children on the ground.”

However, the same evaluation later notes that UNICEF’s Core Commitments for

---

Children (CCCs) in Humanitarian Action, adopted later in 2010, reflect a commitment to ensuring in the future that assessment-based data is available on a humanitarian situation.\textsuperscript{23}

While less likely to be suited to the needs of the target groups, humanitarian evaluations were more likely to have a higher rating with respect to partnership, likely because of UNICEF’s delivery of humanitarian action through non-governmental organizations (NGOs), and as likely to be aligned with national goals as programming in other MTSP focus areas. The global evaluation of UNICEF’s education programming in emergencies highlighted the relevance of this programming, emphasizing the overlap between the countries in which UNICEF is programming and the lists of fragile and conflict-affected countries developed by other organizations. It notes that the “strong focus on capacity-building of government personnel and the heavy investment of funds and resources in delivery of learning materials supports the case for relevance at country and local levels.”\textsuperscript{24}

Partnership was a sub-criterion for assessing the relevance of UNICEF-supported programming and the evaluations identified a wide range of partners involved in programming, including:

- National governments and institutions – a primary UNICEF partner;
- Civil society, as represented by non-governmental and community-based organizations;
- Local government and local authorities;
- Other UN agencies and international organizations; and,
- Bilateral donors and their institutions.

The benefits of partnerships are identified in some evaluations including:

- Strengthening the capacity and engagement of national and local organizations, including improved national and local ownership and strengthened technical capacity and skills for advocacy and social mobilization;
- Improved planning, implementation, coordination and monitoring of activities, including the development of common approaches and shared standards and monitoring mechanisms;

\textsuperscript{23} Ibid., p. 34.

\textsuperscript{24} Progress Evaluation of UNICEF’s Education in Emergencies and Post-Crisis Transition Programme (EEPCT), Evaluation Office, UNICEF, December 2010, p. 70.

OECD DAC Network on Development Evaluation
- Reduced costs through sharing resources and engaging more partners to support activities (Box 4); and,
- Sharing information and best practices.

**Box 4: Burden-sharing by UNICEF Partners**

“[The campaign] was effective because of the establishment of social mobilization committees that included health and other governmental sectors [twelve ministries and institutions, including universities, training institutions etc.] ... NGOs and partners [nineteen national and international partners]. The contribution of the representatives of these sectors increased the strength of the social mobilization campaign. Different actors propose activities to be implemented during the campaign or make a material contribution to facilitate the campaign.” [Translation by authors]


Specific examples of relevance from the evaluation reports include:

- Development of innovative practices, based on international studies, for alternatives to formal education for children and girls in rural villages (Burkina Faso);
- Activities in the education sector after the earthquake and tsunami were developed based on a needs assessment and coordinated by the Taskforce on Education in Emergency Situations, chaired by the Ministry of Education and Human Resources Development with representation from the development partners (Western Solomon Islands);
- Situation analysis for orphans and other vulnerable children that identified problems of malnutrition, morbidity, lack of access to basic education, psychological trauma and abuse (Djibouti);
- Maintenance of an existing network of partners and creation of new partnerships at the local level, including civil society, for water, sanitation and hygiene (WASH) activities (Ethiopia);
- Strong program management and collaboration, including regular consultation and involvement of line ministries and other stakeholders in developing the program, pre-implementation briefs and orientation provided to partners, using a participatory approach in developing program cooperation agreements, regular monitoring and feedback and timely implementation of management decisions and follow-up (Liberia);
• UNICEF playing a leadership role in establishing national working relationships for disaster risk reduction in vulnerable communities and institutions, while still respecting the leadership of local authorities (Central Asia and South Caucasus).

3.1.3 Contributing Factors
A number of factors were identified as enhancing the relevance of UNICEF-supported programming. Most evaluations reflect that UNICEF-supported programming was aligned with national priorities, as expressed in national plans, legal or development frameworks or national programs, policies and action plans. In addition to the alignment with government priorities, UNICEF’s programming was designed to support the implementation of public policies and support programming and capacity building in national governments or institutions (Box 5). Most also noted the extent to which partnerships contribute to relevance.

Box 5: Alignment of UNICEF-supported programming
“The objectives of UNICEF’s Education Programme are highly congruent with international priorities. UNICEF’s education programming in Bangladesh reflects the principles of the Convention on the Rights of the Child (CRC) … [is] aligned with MDGs to achieve universal primary education (MDG 2) and promote gender equality (MDG 3) and article 10 of the Convention on the Elimination of Discrimination against Women (CEDAW). The programme reflects the aims of the Education for All movement and the Dakar Framework for Action…UNICEF’s Education Programme is well aligned with national priorities as expressed in GoB policies, acts, and strategies. …UNICEF is very responsive to country needs, thanks to its close relationship with the government and its participatory approaches and grassroots consultations. The Education Programme is strongly aligned with UNICEF’s corporate priorities for education. Its objectives … are fully aligned with the MTSP 2006-2013 Keys Result Areas under Focus Area 2 (Basic Education and Gender Equality), and with specific aspects of the Key Result Areas under Focus Areas 1 (Young Child Survival and Development) and 3 (HIV-AIDS and Children).”


Some evaluations noted that:
• Programming was aligned with other international priorities, as reflected in UNICEF’s mandate, the UNDAF and national commitments to international agreements, such as the Millennium Development Goals, Convention on the Rights of the Child, the Hyogo Framework for Action and Beijing Declaration; and,

25 The contributing factors are presented in relative order of the frequency with which they were found in the evaluation reports and, as noted in Box 3, were classified as “most” (found in over three-quarters of the relevant evaluations), “many” (between half and three-quarters), “some” (between one-fifth and half) and a “few” (less than one-fifth).
Analysis and research contributed to identifying target group needs and improving program design. Of these, a few noted the contribution of existing research and information, such as international studies, previous evaluations of UNICEF programming and international best practices. However, slightly fewer evaluations noted the use of situation analyses and other assessment tools, including baseline studies, in program design. These were more likely to be identified in evaluations of programming in LDCs.

Although UNICEF-supported programming is seen as being highly relevant, there are a number of factors that detract from the relevance of the programming. These relate primarily to weaknesses in the same areas as the enhancing factors. Some evaluations noted gaps in UNICEF’s partnerships, notably with national government (often because government is not taking adequate ownership) and other UN agencies (Box 6).

**Box 6: Positive and Negative Aspects of Partnership and Coordination**

“Working closely with the RC/HC [Resident Coordinator/Humanitarian Coordinator], UNICEF coordinated its response including needs assessments and external communication with the UNCT [UN Country Team]. In Georgia, coordination was influenced by additional factors, such as the arrival of many external NGOs and donor-led partners who had no previous experience of Georgia or previous relationships with existing humanitarian partners, and a government and UNCT that had had no previous experience of the new humanitarian coordinating mechanism namely the cluster approach. In addition there were several kinds of conflicting pressures by both donors and Government agencies to adopt certain approaches to relief and other interventions that were not all agreed on by the current actors.

In the WASH sector, ... the coordination was focused primarily on the operational aspects of provision of WASH facilities in collective centres and tents. The cluster found it difficult to engage with government, donors and the shelter cluster to ensure that minimum standards in WASH were followed in building settlements, with the result that many of the settlements which the government has now come up with do not have even the minimum habitable infrastructure on WASH.”

UNICEF’s Response to the Georgia Crisis: Real Time Evaluation, March 2009, p. 30-1

A few evaluations noted:

- Either the lack of, or weaknesses in, the analysis of target group needs, such as poor situation analyses, reliance on poor data and little coordination among stakeholders in conducting the needs assessments. This was somewhat more likely to be identified in evaluations of programming in LDCs than MICs;

- Challenges in ensuring effective partnerships, including the number of agencies engaged in a sector, ineffective coordination and weak capacity in partner organizations (including national and international partners);

- The fact that partnerships are not stable over the life of programming – they can both improve (for example, through strengthening coordination) and decline over
time (for example, through the loss of partners; and,

- Challenges to partnership because of frequent staff changes and UNICEF’s lack of expertise in a specific sector (for example, WASH).

### 3.2 Objectives Achievement

#### 3.2.1 Coverage

There is generally strong coverage of sub-criteria relating to objectives achievement in the evaluations (Figure 9). This holds true for objectives achievement (3.1), benefits for target group members (3.2), and for program results reaching substantial numbers of beneficiaries (3.3). Only for the Sub-criterion of positive changes in national policies and programs (3.4) does coverage decline to the moderate level, with 43 evaluations including relevant findings.

The evaluations dealing with programming in Basic Education and Gender Equality (10 of 11 evaluations); Child Protection (5 of 6 evaluations); and evaluations of Humanitarian programs (10 of 13 evaluations) do tend to address UNICEF’s success in influencing national policies and programs quite often. Only two of seven evaluations in the sample from the focus area Policy Advocacy and Partnership included findings on Sub-criterion 3.4.

One explanation for only moderate levels of coverage of the extent of programs’ influence on changes in national policies may be the close relationship between UNICEF and national policy makers in some countries: extending over a number of policy making cycles. At the time of the evaluations, there may have been little difference between UNICEF priorities and national policies, leaving little need for UNICEF to influence changes or to evaluate those changes.
3.2.2 Key Findings

The review findings are largely positive (Figure 10). A full 77% of the evaluations reviewed reported satisfactory or better findings regarding the extent that UNICEF supported programs achieve their stated objectives (3.1). This does not mean that UNICEF programs were found to achieve 77% of their objectives across the board. Rather it means that the UNICEF programs were reported as achieving better than half their objectives (including the most important ones) in 77% of the evaluations reviewed.

Positive results were even stronger for the programs’ ability to provide positive benefits for the target population (3.2) with 86% of evaluations reporting satisfactory or better findings. They were somewhat less positive for the scale of program benefits measured in terms of the number of program beneficiaries (3.3) with 72% of evaluations reporting satisfactory or better results.

Finally, programs were rated satisfactory or better for their ability to support positive changes in national policies and programs (3.4) in 77% of the evaluations reporting on this sub-criterion.
Of the four sub-criteria relating to objectives achievement, the last three showed no notable difference in the pattern of findings when evaluations of programs in LDCs were compared to those of programming in MICs.

On the other hand, for the key sub-criterion of whether programs achieve their stated objectives (3.1), evaluations of programs in MICs reported satisfactory or better findings somewhat more often than for programs in LDCs (Annex 6).

This does not seem counterintuitive since the program implementation capacity of UNICEF’s partner agencies in MICs might in general be higher than in LDCs. The pattern of more positive performance for programs in MICs does not hold, however, for all the criteria, as can be seen from the results reported for gender equality (Section 3.3).

**Types of Objectives Achieved by UNICEF-Supported Programs**

It is interesting to consider what types of objectives are being achieved by UNICEF-supported programs according to the evaluations. It is clear from the pattern of programming represented by UNICEF’s MTSP focus areas and the sub-sectors which they encompass, that a wide variety of program outputs and outcomes are included in the objectives of UNICEF programming.

Providing an exhaustive portrait of the positive achievements of UNICEF-supported programs and how they contribute to the goals of the MTSP is beyond the scope of this review. Instead this section presents a small sampling of the specific achievements noted in the evaluations. These objectives were achieved in partnership with governments,
NGOs and local community members. The evaluations do not attempt to attribute their achievement solely to UNICEF support. The achievements reported in the evaluations include:

- Changes in knowledge, understanding and behaviour towards the risk of anti-personnel mines (Colombia);
- Establishment and promotion of good practices in WASH through technical guidelines and improvements in national clean water access through a shift toward community managed systems (Sri Lanka);
- Improvements in the treatment of children and adolescents by police services in tsunami affected regions of Indonesia (Box 7);

**Box 7: Improvements in Police Services in Indonesia**

“By 2007, the creation of children’s desks in every district in Indonesia was agreed to by the national police. A component on child-friendly procedures was also incorporated into the police training curriculum. By 2008, 22 children’s desks had been re-established in all districts of NAD; one child courtroom established in Banda Aceh district; instructions on diversion [of youth to the restorative justice system] adopted by the police; standard procedures and guidelines for restorative justice for police officers developed; and a case management database was under development.”


- Improvements in the treatment and opportunities for disabled children in a family and community setting (Mexico);
- Improvements in sanitation and hygiene through construction of latrines and the use of a Community Led Total Sanitation Approach (Mauritania);
- Re-integration of children of families expelled from Tanzania into the education system using a Child Friendly Schools approach (Burundi);
- Rapid provision of immediate relief assistance following the earthquake in Haiti by restoring water supply to Port-au-Prince and through provision of non-food relief items (Haiti);
- Increased real household consumption and food expenditures and dietary diversity for poorer households (with a positive impact on secondary school enrolment) as a result of a cash transfer program for Orphans and Vulnerable Children (Kenya);
- Increased adolescent and youth participation in project implementation as part of the UNICEF supported Advocacy and Social Mobilization Programme (Cambodia):
Improvement in the social status and self-esteem of women and children participants in radio programming and in live debates with local authorities and development partners through a community radio listening groups project (Ethiopia);

Increasing equitable access to health care through incorporation of an Integrated Management of Childhood Illnesses approach and through changes in the knowledge and practices of caregivers (Moldova);

Improvements to household clean water supply and sanitation services (Ethiopia) (Box 8); and,

**Box 8: Contributing to Improved Sanitation and Hygiene in Ethiopia**

“The survey results showed that 93% of the households interviewed had access to an improved water source which is above the stated level of achievement while access to a latrine was 74% which is slightly less…The improvement in service levels achieved in the target woredas [administrative districts in Ethiopia managed by local governments] can be regarded as one of the measures of impact of the project. Additionally, woreda staff report a genuine change in attitude and commitment of the community as a demonstrable impact of the project. The role of WASH committees in ensuring sustained operation of water schemes and monitoring of the sanitation condition of their villages was also considered important. In general there was considerable anecdotal support behind the project’s contribution to minimize community susceptibility to diseases emanating from poor hygiene and sanitation conditions.”

*Ethiopia: Mid-Term Evaluation of EU/UNICEF Supported WASH Programme, 2010*, p. 21 and 32

Swifely responding to the food security needs of children under two and pregnant and breastfeeding women in an emergency setting through the provision of cash transfers (Niger).

### 3.2.3 Contributing Factors

Factors which made a positive contribution to the achievement of program objectives are presented in this sub-section.

Many evaluations pointed to the following factors:

- High quality program designs encompassing child friendly elements and clear objectives that were often multi-sectoral in their approach and fully integrated into national sector policies. This included programs which were developed jointly with government partners;

- The positive effect of UNICEF’s active role in developing and influencing national policy at the sector level; and,

- An appropriate scale and size of programming in resource terms and in scope.
relative to the problem being addressed as an important factor in program success.

Some evaluations pointed to the following positive factors:

- Appropriately targeted capacity development efforts and a strong emphasis on capacity development for key stakeholders;
- Effective lobbying and advocacy by UNICEF as a factor in ensuring that key issues such as child protection, gender-based violence, girls’ education, youth participation, and child rights were given priority in program design and implementation;
- A link between an inclusive process of program design and implementation which drew in all important stakeholders (UN, government agencies, Civil Society Organizations (CSOs), community leaders, religious leaders, and girls and women) and successful program results (Box 9);

**Box 9: Strengthened Programming through Participatory Processes: United Nations Girls Education Initiative in Uganda**

“Indeed, UNGEI Uganda provides a model for multi-stakeholder ownership and management of girls’ education, whereby the school community has been firmly placed “in the driver’s seat”. The school communities are directly consulted on matters relating to their priority needs and perceptions about the quality of education service delivery. The findings from such consultative efforts are factored into the work plans and are used to inform resource allocation decision-making. As a result, there is now increased prioritization of school community-specific interventions and raised grassroots’ consciousness.”


- The effective targeting of marginalized or vulnerable groups (including targeting of girls and women). This factor was identified more often in evaluations of programs in LDCs than it was for those in MICs; and,
- The high technical quality of UNICEF-supported technical and methodological guidelines as an important element in promoting good practices in service design and delivery. This factor was also noted more often in evaluations of programs in LDCs than for MICs.

A few evaluations highlighted each of the following positive contributing factors:

- An early or rapid response by UNICEF as a key factor in preventing exploitation and abuse or in providing key inputs during humanitarian emergencies;
- In humanitarian action, the role of appropriate attention to the repair, rehabilitation or construction of critical infrastructure, especially in health and education;
• Effective monitoring and information gathering as a basis for ongoing follow up helped to improve program understanding and strengthening of good practice in the program; and,

• The presence of a dedicated and well qualified project implementation team (including UNICEF country office staff, hired project staff and implementing partner personnel) as an important factor in project success.

Global evaluations echoed some of the same factors for success in securing positive program results and achieving objectives. The synthesis report on Protecting Children from Violence, for example, highlighted advocacy and the incorporation of child protection into national and decentralized planning processes as an essential tool for strengthening child protection systems.26

The evaluations also provided a diagnosis of some factors which negatively affect the ability of UNICEF-supported programs to achieve their stated objectives.

Many evaluation reports referred to weakness in program design as the crucial factor inhibiting program success. These included lack of a results orientation, absence of linkages or a total absence of a causal chain, over-ambitious program goals, poorly targeted programming, failure to address the rights of target group members, an overly fragmented approach to program delivery, and poor technical design of program inputs. Poor program design was more often noted in evaluations of programming in LDCs than in MICs.

Some evaluations pointed out that program budgets were inadequate and that outputs were insufficient to address the identified problem and to reach the coverage targets set for the program.

A few evaluations highlighted the following negative factors influencing objectives achievement.

• Weaknesses in the capacity development component of programs including a general lack of attention to capacity development and training needs, inadequate investment, and technical weaknesses in training plans and packages;

• Problems with poorly trained or qualified program staff (facilitators, teachers, animators, service delivery personnel) sometimes associated with low pay and high turnover rates. Interestingly, all of these references were to programs in LDCs;

• Weaknesses in program monitoring as an impediment to learning and to program

---

improvement over time, and hence to objectives achievement;

- Lack of a gender perspective or a failure to mainstream gender equality impedes program success; and,
- The late arrival or insufficient supply of planned programme inputs.

3.3 Cross-Cutting Themes

3.3.1 Coverage

Compared to Relevance and Objectives Achievement the sub-criteria relating to Cross-cutting Themes present a very different picture in terms of coverage in the UNICEF evaluations (Figure 11). Sub-criterion 3.1 on the extent UNICEF supported activities effectively address gender equality received a moderate level of coverage with 42 of the 62 evaluations containing relevant findings. Coverage of Sub-criterion 3.2 on the environmental sustainability of changes resulting from UNICEF supported programs was weak. It was addressed in only 15 evaluation reports.

Figure 11: Number of Evaluations Addressing Sub-criteria for Cross-Cutting Themes

Sub-Criteria 3.1 Gender Equality

It is remarkable that almost one-third of the evaluations in the sample did not address the question of gender equality. It is important to note that in order for the review to rate this sub-criterion, evaluations had to identify the extent to which the program or project in question had incorporated gender equality objectives and whether they had been
achieved. Gender equality could also be addressed in an evaluation report if it assessed the program’s success in mainstreaming gender equality. On the other hand, if an evaluation simply reported the portion of girls or women receiving benefits compared to boys or men, not addressed was the rating provided for this criteria because there was no reference to the program’s success in addressing equality.

It is also worthwhile considering which evaluation reports in the sample tended to address gender equality more than others. Evaluations of UNICEF support to programs in the MTSP focus area of Basic Education and Gender Equality tended to address gender equality with ten of eleven reports covering the issue. Similarly, eleven of thirteen evaluations of humanitarian action adequately covered gender equality.

Lower levels of coverage were found in evaluations of programming in the Young Child Survival and Development focus area with only 13 of 21 evaluations addressing the subject adequately. This may reflect the operational nature of much programming in this area, which includes support to Water, Sanitation and Hygiene (WASH) programs. On the other hand, programs in child health could conceivably often face issues of differential access for boys and girls.

Finally, some evaluation reports pointed out that UNICEF and partner staff felt that gender equality was implied in programs which implemented a results-based approach or which focused on child protection and/or child rights.

The recent global evaluation of the application of the Human Rights Based-Approach to programming notes that UNICEF responded to an earlier global evaluation of its gender equality policy by re-emphasizing its organizational commitment:

> The evaluation found that the strategies advocated by the original policy [mainstreaming and gender-specific programming activities] remain sound, and these have been retained. However, it indicated that the policy required updating to respond to new programme priorities. Such priorities included the commitment to work more explicitly with men and boys as both agents and beneficiaries of gender equality, and to improve the priority and resourcing given to gender-equality programming by the organization and the grounding of its actions in CEDAW together with the CRC.27

Given the strength of UNICEF’s strategic commitment to gender equality and its re-

---

emphasis in the 2010 policy, it is surprising that one-third of the evaluations reviewed fail to address gender equality.

**Environmental Sustainability**

Given the very weak level of coverage of environmental sustainability, with just 15 evaluations addressing the sub-criteria, the review is not able to report on the effectiveness of UNICEF-supported programs in this area.

It is worth noting that evaluations of Young Child Survival and Development programming do sometimes address environmental sustainability (eight of 21 evaluations). This is largely due to the presence of WASH programs in this focus area with consequent attention paid to the environmental effects of investments in water and sanitation infrastructure.

Similarly, evaluations of humanitarian action also sometimes addressed environmental sustainability (six of 13 evaluations). This is not surprising since emergency programs often include a component focused on the rehabilitation of water, sanitation, education and health infrastructure.

The Programme Division of UNICEF reports that programmatic attention to climate change adaptation grew in 2012, especially in relation to WASH programs. Given this trend and the activation (in 2012) of an inter-divisional working group on climate change adaptation, it would be reasonable to expect increased attention to issues of environmental sustainability in evaluations carried out by UNICEF in the future.

### 3.3.2 Key Findings in Gender Equality

The evaluations present a challenging picture of the effectiveness of UNICEF-supported programs in addressing gender equality. Less than half (47%) of the 42 evaluations that addressed gender equality contained findings of satisfactory or better for this sub-criterion (Figure 12).

Considering that 22 evaluations did not address gender equality, another way of expressing the result would be to note that only one-third (32%) of the 62 evaluations reviewed were able to demonstrate that UNICEF-supported programs effectively address gender equality.

---

Surprisingly perhaps, results for effectively supporting gender equality were somewhat more positive when reported in evaluations of programs in LDCs than those in MICs (Annex 6).

At the same time, results for Sub-criterion 3.1 were notably more negative for humanitarian action than for the samples of evaluations taken as a whole (Annex 6). So while evaluations of humanitarian action were more likely to address gender equality, they reported more negative findings than other evaluations.

However, the evaluations do point to some important successes for UNICEF in supporting programs that effectively address gender equality. Where programs are effective, they have been reported to achieve many different types of successful results in gender equality. Just a few examples include:

- Extension of psychosocial programs to include child abuse and to address behavioural problems within families that leave girls at risk (Sri Lanka);
- Improving safety in school attendance for girls in Child Friendly Schools programs (Burundi);
- Improved access to basic education for girls (Burkina Faso);
- Promotion of the adoption of national laws on child-trafficking and the elimination of domestic violence, including Gender-Based-Violence (Indonesia);
- Improved access for women and adolescent mothers to peri-natal care (Senegal);
- Improved access for women and girls to safe water where women have been promoted as active participants and managers of water facilities and water management committees (Ethiopia);
- Promotion of women to key national policy making positions relating to early

**Figure 12: Findings for Effectiveness in Supporting Gender Equality**

![3.1 Effectively Address gender equality (n=42)](image)

<table>
<thead>
<tr>
<th>Category</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>(4) Highly Satisfactory</td>
<td>7%</td>
</tr>
<tr>
<td>(3) Satisfactory</td>
<td>40%</td>
</tr>
<tr>
<td>(2) Unsatisfactory</td>
<td>38%</td>
</tr>
<tr>
<td>(1) Highly Unsatisfactory</td>
<td>14%</td>
</tr>
</tbody>
</table>
childhood development (Ghana); and,

- Improved treatment of girls and women by police services as a result of effective training and sensitization of police officers (India) (Box 10).

**Box 10: Gender Sensitisation and People Friendly Police (GSPP) Initiative in Karnataka**

“Definite changes were seen in the way women and children are treated at all the police stations visited. This was also seen in instances where we could see women interacting with the police at the station. This was also echoed through our discussions with NGOs and community members….Over and above all these, the GSPP training was seen as an important factor for bringing about this change. This was expressed during discussions and in the questionnaire survey.

_Evaluation of the Gender Sensitization and People Friendly Police Initiative: Karnataka, 2011 p. 45_

The global evaluations reviewed qualitatively also point to a variety of positive results in gender equality for UNICEF supported programmes. The global evaluation of Child Friendly Schools Programming, for example, found that:

Students mostly feel that their schools provide female and male students with equal access to opportunities: about three quarters or more in each country responded ‘mostly true’ or ‘very true’ to the statement ‘Both boys and girls have equal opportunities to succeed at this school’.  

It is worth noting that UNICEF has recognized issues relating to gender equality and has been engaged in efforts to improve effectiveness in the area. In response to the 2008 Gender Evaluation, the agency launched the Strategic Priority Action Plan for Gender Equality Action Plan 2010-2012 with eight areas of organizational action identified. UNICEF also reports its intent to take visible action on gender equality in the new strategic plan.

### 3.3.3 Contributing Factors

Given the generally negative findings for gender equality, it is not surprising that reviewers noted more negative than positive contributing factors. Nonetheless, the evaluations discussed and presented a number of contributing factors.

The factor identified most often as contributing to positive results in gender equality was the inclusion of gender specific objectives and targets, including a focus on women’s rights and (for a few evaluations) effective mainstreaming of gender equality beyond a single thematic area (such as Child Protection).

Some evaluations pointed to the active promotion of women as leaders and as members

of water and other community management committees. They also pointed to the promotion of women to key national policy making positions.

A few evaluations noted each of the following factors contributing to positive results in gender equality:

- A program design which included a good diagnosis of the gender situation and barriers to gender equality;

- High quality and consistent efforts to develop gender-sensitized capacity for service delivery among staff of CSOs, Ministries of Health, public health services and police and judicial services could be effective in support of gender equality; and,

- Effective policy dialogue by UNICEF with central and local government officials as a factor contributing to positive results in gender equality.

The reviewed evaluations identified a wide range of factors inhibiting achievements in gender equality.

Most evaluations highlighted the fact that programs lacked any specific objectives regarding gender equality and/or they simply lacked any gender perspective in their design and implementation. For a few evaluations, this took the form of treating gender a completely “neutral” issue in program design, with no distinction made between boys and girls or men and women.

Some evaluations pointed to:

- The absence of gender disaggregated data, either as a baseline for diagnosing challenges in gender equality or as information for use during program implementation and monitoring, as a factor impeding progress on gender equality. This included the use of gender disaggregated data for vulnerability mapping;

- Failure to promote increased inclusion of women in leadership and decision making roles impeded results in promoting gender equality. A few pointed to unequal treatment of program staff with men in paid leadership positions and women relegated to either underpaid or volunteer staff; and,

- A program’s failure to recognize and promote women’s capacity in the technical areas of programming in sectors such as Child Protection or WASH.

A few evaluations identified negative factors of:

- A failure to recognize that some technical program component of a program was adversely affecting women or girls (Box 11);
Box 11: Gender-Based Violence and the Response to the Earthquake in Haiti

“...in the early months of the response, UNICEF had only one gender-based violence focal point. (It took a month for the organization to deploy a gender-based violence specialist), and to this day it remains unclear where the issue sits within child protection. This was despite the emergence of rape in the camps as an issue before the end of January 2010....It was reported by UNICEF surge staff that it had taken up to four months to install lighting for toilet areas in some camps using UNFPA materials – a key security precaution that might have helped combat sexual violence.

*Independent Review of UNICEF’s Operational Response to the January 2010 Earthquake in Haiti*

2011, p. 15

- The lack of emphasis on gender issues in training programs and supporting materials;
- Lack of consultation with women and girls about their specific health and protection needs resulting in lack of attention to women’s or girls’ needs and interests; and,
- Lack of integration of male gender issues or sensitization of male community members.

Some global evaluations identified similar issues for UNICEF in effectively addressing gender equality issues. The synthesis report on Protecting Children from Violence, for example, noted a number of challenges, constraints and gaps faced by UNICEF. The first two mentioned are consistent with the factors reported above:

*Gaps:* In some programmes, evaluators noted that gender issues are not addressed adequately in planning and implementation (Bosnia Herzegovina, Thailand). In some evaluations concerns were raised that programmes were not using their programmes to raise awareness around gender and/or stimulate discussion in this regard (India).

*Knowledge Management:* Failure to adequately conduct an in-depth analysis of gender issues at the outset, hindered programmes from effectively enduring gender equity (Nepal, Maldives).  

The global evaluation of Human-Rights Based Approach to UNICEF programming examines the human-rights based approach under five key principles. While all five are relevant to gender-equality, perhaps the most central is the principle of non-discrimination. The global evaluation also identifies data disaggregation and weaknesses in reporting as factors limiting the application of the principle of non-discrimination in

---


**OECD DAC Network on Development Evaluation**
programming by UNICEF country offices:

The operationalization of this principle in UNICEF’s work rests to a large extent on the degree to which data disaggregation is sufficiently carried out across all levels of vulnerability so as to help inform the development of targeted or universal programmes.

The scoring on non-discrimination across the CO [Country Office] Assessment was ‘satisfactory’ to ‘weak’, suggesting that it is applied with mixed results. More specifically, the CO Assessment rated more than one third of COs (15) as ‘weak’ on this principle. The SitAn [Situation Analysis] often reveals who is being excluded in society, who the marginalized groups are, etc. However, SitAns often use generic descriptions for identifying the most vulnerable, and programming documents and annual reports do not provide sufficiently detailed information about the specific efforts COs make to ensure they reach the most vulnerable/marginalized. Moreover, in many cases, documents did not provide sufficient detail to determine whether the most vulnerable have been reached.31

3.4 Sustainability

3.4.1 Coverage

The coverage for two of the three sub-criteria with respect to sustainability was strong (Figure 13). Both Sub-criterion 4.1 “Benefits continuing after program completion” and Sub-criterion 4.2 “Institutional/community capacity for sustainability” were covered in 59 evaluations. The coverage was moderate for Sub-criterion 4.3 “Strengthened enabling environment for development”, as this sub-criterion was addressed in 40 evaluations.

3.4.2 Key Findings

Overall the results with respect to sustainability are mixed (Figure 14). About half the evaluations reflected ratings of satisfactory or better on two sub-criteria – Sub-criterion 4.1 “Benefits continuing or likely to continue” (51%) and Sub-criterion 4.2 “Programs support sustainability through institutional and/or community capacity” (48%). Evaluations of programming in MICs appear to be somewhat more likely to get satisfactory or better ratings for sustainability through institutional and/or community capacity than those from LDCs (Annex 6). The evaluation reports were more positive for Sub-criterion 4.3 “Programs strengthen enabling environment for development”, with 71% reflecting findings of satisfactory or better.
Box 12: Advocacy and Policy Dialogue with Government

Having established a strong advocacy role with the GoSL [Government of Sri Lanka], UNICEF now has the potential to influence policy changes in water quality surveillance and water supply subsidies [in the WASH sector]. ... In child protection, UNICEF has successfully lobbied for changes in practice and attitudes away from institutions and towards safe placement of children with legal guardians. ... Attributing change to any one agency would be dishonest, but UNICEF sizeable interventions over four years, and the encouraging manner in which it has assimilated lessons from these interventions, have given it a unique opportunity to guide and influence national and sub-national government.


- Advocacy and policy dialogue with government (Box 12);
- Capacity building for both individuals (in terms of staff training for planning, management and monitoring) and institutional building (through the development of training institutions, development of procedures, norms and standards, and the encouragement of cross-sectoral collaboration mechanisms); and,
- To a somewhat lesser extent, development of community capacity and encouragement of community engagement.

Specific examples from the evaluations include:
- Advocacy with the national government to influence policy changes in water
quality surveillance and water supply subsidies (Sri Lanka);

- Building technical capacity of national, provincial and local level staff on child injury prevention and forming a wide range of alliances with various government ministries and international partners, through project implementation (China);

- Development of a protective environment strategy to promote the social welfare and juvenile systems and to expand programmatic coverage to include children and youth in post-conflict areas following the tsunami (Indonesia);

- Use of a linking relief, rehabilitation and development approach to support the needs and reintegration of persons expelled from Tanzania (Burundi); and,

- Use of existing local resources through a network of partner organizations to implement mine risk education activities (Nepal).

It was noted that the results for Sub-criterion 4.3 are more likely to be positive than those for other sub-criteria. It is suggested that this is because efforts to improve the enabling environment do not, in themselves, have to be on-going activities. For example, support for the development of national legislation or policies is a time-limited activity, which would be expected to have a long-term impact on the sector. As a result, it may be easier to achieve satisfactory findings for this sub-criterion than for other sub-criteria, which require evidence of the long-term impacts. This is reflected in some of the achievements with respect to strengthening the enabling environment:

- Creating a supportive environment for early childhood development helped to promote the establishment of formal program for training early childhood educators, even though only a small proportion of teachers are formally trained (Ghana);

- Enhanced processes for participation by civil society in development-related issues (Sudan); and,

- Setting up of a National Learning Platform, which included project consortium members, to facilitate the review of WASH sector policies and strategies. This led to the adoption of new sector coordination mechanisms (Zimbabwe).

3.4.3 Contributing Factors

A number of factors identified in the evaluations were linked to enhancing the sustainability of UNICEF programming. Some evaluations reflected:

- A link between sustainability and programming that included strong national government engagement. This included programs that were adequately integrated
into national programs (for example, were managed by national staff), included strong partnerships with national ministries and institutions, and had components of capacity building for national authorities. Evaluations of programming in MICs were somewhat more likely to demonstrate this factor; and,

- A link between sustainability and programming that included strong local level and community engagement, including programs that strengthened local and community structures and incorporated training for local authorities and community workers.

A few evaluations noted that:

- It was the design and implementation of programs that contributed to their sustainability. Designs that contributed to sustainability included a sustainability strategy based on clear needs, focused on developing user responsibility for the program and included an exit strategy. Similarly, implementation that went well and achieved program objectives contributed to sustainability;

- Sustainability was improved when adequate financial resources were available, either from national governments or other donors; and,

- Decentralization of government structures contributed to sustainability.

The major factors detracting from sustainability were, for the most part, the failure to achieve the factors that enhance sustainability. Some evaluations noted:

- A failure to plan for sustainability, in that the programs did not have exit strategies and lacked a focus on sustainability at the institutional level. They often did not include sufficient capacity building activities, or included capacity building that focused on technical aspects, rather than the cultural changes necessary to ensure sustainability (Box 13). They did not give sufficient attention to strengthening the enabling environment for sustainability;

- Weaknesses in program design and implementation limited sustainability. This included programs that did not have a program strategy or were not sufficiently integrated into national programs. The sustainability of some programs was

---

**Box 13: Lack of strategy for sustainability**

“There was insufficient explicit attention to sustainability of results in the CPC 5 and CPC 6 [UNICEF Country Program for Children 5 and 6] programs. While capacity building was central to both programs, the absence of a common understanding, clear strategy, and systematic approach to capacity building within the UNICEF Philippines Country Office (PCO) contributed to its mixed performance in supporting sustainable capacity building results.”

reduced by the failure to achieve program objectives or by inconsistencies in the implementation and success of programming across communities. It also included programs that were too short, not reflecting that it takes time to achieve sustainability;

- Financial limitations to sustainability, including, most commonly, an ongoing dependence on donor support and government inability to take on continued program funding, but also included the inability of communities and users to pay for services. Financial limitation were more likely to be identified for programming in LDCs; and,

- The lack of institutional capacity at the national level (for example, lack of national policies and government funding, weaknesses in coordinating structures, and lack of government monitoring) and at the community level.

In addition, a few evaluations noted that, in spite of efforts at capacity building, a challenge to sustainability lies in the loss of previously developed capacity – for example, the faltering of stakeholder engagement, the disbanding of community committees and the dropout of trained resources.

A few evaluations of humanitarian action reflected that weaknesses in program design detracted from sustainability because of the limited linking of relief and development (Box 14). Linking relief and development is reportedly being addressed by UNICEF in its upcoming new strategic plan, which will highlight UNICEF’s commitments to building resilience and addressing vulnerability and fragility through its development programming.\(^{32}\)

\^ {32} See also 2012 UNICEF Humanitarian Action for Children, UNICEF, January 2012, p. 12-14
Box 14: Lack of linking of relief and development

“In the health sector, a successful emergency phase was not matched by strategic multi-year planning in the recovery phase. More attention should have been paid to the development of district- and provincial-level capacity for planning and administration. Structural weaknesses in government, apparent before the tsunami, persist; UNICEF’s preoccupation with bricks and mortar has prevented appropriate attention to this. Nevertheless, UNICEF has helped to build institutional capacity through training of health care workers, social workers and volunteers…”


3.5 Efficiency

3.5.1 Coverage

Coverage levels for the three sub-criteria of efficiency range from strong for Sub-criterion 5.1 on the cost and resource efficiency of programming to moderate for Sub-criteria 5.2 and 5.3 (Figure 15). The last two deal with the timeliness of program implementation and the effect of UNICEF’s administrative systems and procedures on the efficiency of program delivery.

It is important to point out that the fairly strong level of coverage of cost efficiency (5.1) does not mean that UNICEF programs are regularly subjected to cost effectiveness measurement in the evaluations. In fact, 28 evaluations in the sample did not include any information on the overall cost of the program.33 What is meant by coverage of cost and resource efficiency is that the evaluations made some judgment on program efforts to control the resource requirements and/or the unit cost of program inputs or outputs. Nonetheless, the fact is that 48 of 62 evaluations in the sample (77%), made some effort to assess the cost efficiency of the programs under review.

---

33 The review process required the raters to identify the overall costs of the programs. Information was found in only 34 evaluations and the nature of the information presented was inconsistent across evaluations – at sometimes reflecting the overall costs, at some only UNICEF’s costs; at times reflecting the budget, at times the actual expenditures. As a result, the team was not able to provide any information on the overall value of the programming being covered in the evaluation sample.
The moderate level of coverage of sub-criteria on timeliness (5.2) and administrative systems and procedures (5.3) is important. The review team suspects that these two criteria tend not to be addressed routinely or be included systematically in evaluation terms of reference. They may be more likely to be included in evaluation reports when problems in timely delivery of program inputs or problems with administrative systems and procedures are encountered during field level evaluations. The information may then be presented to explain limitations in the achievement of program objectives.

For Sub-criterion 5.1 on cost and resource efficiency there is no particular pattern in terms of which MTSP focus areas tend to address the issue. Evaluations in all focus areas (and Humanitarian program evaluations) examine cost efficiency in about three-quarters of evaluations in the sample.

The pattern is similar for timeliness of program delivery (5.2) with the exception that only one of six evaluations in Child Protection addressed this sub-criterion. For Sub-Criterion 5.3 on the efficiency of implementation systems and procedures, coverage levels are low for all focus areas except for Young Child Survival and Development (addressed in 14 of 21 evaluations) and for humanitarian action (addressed in nine of 13 evaluations).

For both these two focus areas, on-time delivery and installation of elements of infrastructure are often crucial to achieving program objectives, with the result that evaluations often report findings relating to the effectiveness of program implementation systems, particularly systems of procurement and supply.
3.5.2 Key Findings

The findings for efficiency are mixed and require careful interpretation. The fact that 62% of evaluations reported satisfactory or better findings on the cost efficiency of programs represents a reasonably positive result for Sub-criterion 5.1 (Figure 16). Also notable is the fact that no evaluations reported highly unsatisfactory results on cost efficiency.

As already noted, there may be a tendency to under-report findings with respect to timeliness and the efficiency of administrative systems when no problems are encountered at field level. Nonetheless, it is a concern that 49% of the evaluations that addressed timeliness reported findings of unsatisfactory or worse. Also worrying is the fact that 18% reported a finding of highly unsatisfactory.

Similarly, the findings for Sub-criterion 5.3 on the efficiency of administrative systems and procedures are largely negative with 63% unsatisfactory or worse and 17% highly unsatisfactory.

Figure 16: Findings on Efficiency

Interestingly, this is an area where results for evaluations of programs in LDCs are somewhat more positive than for MICs. In particular, results on the timeliness of program implementation are notably more positive for programs in LDCs (Annex 6).
The evaluation reports contain some interesting examples of the ways UNICEF-supported programs have tried to reduce unit program costs and increase the efficiency of resource use by:

- Entering into partnerships with other UN agencies and NGOs to reduce transport costs (Burundi);
- Emphasizing low cost, sustainable and scalable solutions such as Community-Led Total Sanitation (Ethiopia and Cambodia);
- Using pre-existing resources and pre-established delivery channels for program implementation (Mozambique);
- Providing training in situ, taking the skills to the participants instead of transporting them to urban centers (Kenya); and,
- Constantly assessing alternative program delivery approaches and making necessary modifications to reduce costs (Uganda).

3.5.3 Contributing Factors

Because very different factors contribute to results regarding cost efficiency, when compared to timeliness and the efficiency of administrative systems, they are discussed separately in this section.

*Sub-Criteria 5.1: Cost/Resource Efficiency*

Positive factors promoting the cost efficiency of programs cited by some evaluations included the following:

- Effective monitoring systems allowed for regular tracking of costs and continuous improvements to efficiency through lower delivery costs or lower unit prices for outputs;
- The use of low-cost approaches to service delivery such as Community Led Total Sanitation or the use of insecticide-treated bed nets;
- The use of low cost arrangements for training and capacity development such as the use of informal trainers and training delivered on-site; and,
- The strength of strong financial planning during program start-up and continued use of strong financial and cash-flow management during the program.
A few evaluations pointed to:

- Competitive tendering for construction services and/or supply contracts allowing bidding by for-profit firms, NGOs and CSOs which led to lower unit costs (Box 15);

**Box 15: Open bidding for well drilling in Ethiopia**

“Among donor financed WASH programs however, it is the UNICEF program alone that allows public enterprises, NGOs and private contractors to participate equally in the bidding process. Particularly in the SNNPR [Southern Nations and Nationalities’ People’s Region] region, church-based organizations and other NGOs are awarded contracts to drill UNICEF financed wells and the regional water resource bureau also possesses [a] rig donated by UNICEF which is mainly used to drill UNICEF financed wells. This has contributed to keeping the cost of drilling down.”

*Mid-Term Evaluation of EU/UNICEF Supported WASH Program, Ethiopia, 2010. p. 14*

- Community involvement in service delivery contributing to savings over comparable salaries in the formal sector; and,

- Reduced transportation costs through partnership with other agencies.

Negative factors contributing to reduced cost and resource efficiency for UNICEF-supported programs were also identified in the review of evaluations:

- By far the most frequently cited factor impeding the cost and resource efficiency of UNICEF-supported programs is the lack of appropriate cost data reported regularly and available on time to allow for a reasonably accurate calculation of service costs. Where this information was available, it was sometimes subject to delays or significant gaps in the data available; and,

- Some evaluations pointed to delays in programs start-up or delays in the construction of facilities and supply of inputs as a factor in raising costs and reducing efficiency.

A very diverse set of other negative factors for cost efficiency were identified and reported in a few evaluations. They include:

- High staffing transaction costs;

- High staff turnover requiring expensive re-training;

- Duplication of effort among agencies contributing to higher costs and procurement of un-needed inputs;

- Procurement and installation of over-designed, expensive equipment or infrastructure (Box 16);
Box 16: Complex and Expensive Sanitation Systems for the Maldives

“The potential cost of running these systems for the islands is high, the human and institutional capacity to manage the complex systems is low, and key policy frameworks are still in development. The sanitation systems built as a response to the tsunami are complex engineering feats that are over-designed and expensive. This new method necessitates cooperative management and financial recovery, which is not yet available on the islands.”


- Unnecessarily high transport costs; and,
- Weaknesses in targeting which reduce cost efficiency by raising the unit cost of reaching each member of the actual target group.

Sub-Criteria 5.2: Timeliness of Program Implementation and 5.3: Effective Systems and Procedures for Program Implementation and Follow-Up

There was a very close relationship between the timeliness of UNICEF-supported programs and the efficiency of UNICEF systems and procedures for program implementation as reported in the evaluations. For that reasons, the factors which supported or impeded positive results for these two sub-criteria tended to be identical. They are presented together in this section.

Positive factors for sub-criteria 5.2 and 5.3 reported in a few evaluations include:

- The strong management and programming capacity of UNICEF country and regional offices which allowed them to respond to program demands rapidly. In the case of humanitarian action, this factor contributed to UNICEF rapidly establishing cluster coordinating structures where it was the lead agency. It also helped UNICEF offices work around difficult administrative procedures;
- Timely coordination among partner organizations (including UNICEF) allowed for expedited decision making in response to rapid changes in the program context;
- Effective program monitoring and regular supervision which allowed for rapid responses to changes in the program; and,
- Effective cash flow management by UNICEF.

Evaluations also identified negative factors which tended to limit the timeliness of UNICEF-supported programs, often related to administrative systems and procedures. Some evaluations pointed to:
• Lengthy delays in procurement by UNICEF (Box 17); and,
• Rigid and cumbersome UNICEF financial systems and procedures for funds disbursement delayed program implementation.

**Box 17: After a Rapid Start, Delays in Emergency Response in Yemen**

“Despite its general lack of preparedness for emergencies and its dominant development orientation, the UNICEF Country Office in Yemen acted relatively quickly in 2009 to mobilise existing staff and supplies for the emergency response. A loan of two million US$ from UNICEF’s internal emergency financing mechanism, the Emergency Programme Fund (EPF) in September 2009, enabled this rapid response. Subsequently, however, it experienced often significant delays in procuring or shipping additional supplies or implementing activities that are mainly due to the fact that the Country Office is relying on its development structures and processes which grant little autonomy to local staff and have slow procurement and supply procedures (cf. section 3.6). These delays undermine UNICEF’s credibility in the eyes of its partners and beneficiaries. As a result, beneficiaries appear less open to accepting important “soft” interventions by UNICEF and its implementing partners, such as hygiene education or awareness raising activities on child protection issues.”

*Real-Time Evaluation of UNICEF’s Response to the Sa’ada Conflict in Northern Yemen, 2010, p.23*

A few evaluations identified:

• Overly ambitious program timeframes eventually led to delays;
• The lack of timely program oversight and weak monitoring systems as factors hindering UNICEF’s ability to respond to program bottlenecks and quickly resolve delays;
• Lack of credible baseline information leading to longer time frames for program development and delayed program startup; and,
• Limited capacity to absorb funding on the part of the implementing agency.

The global evaluations of humanitarian action also identified similar strengths and challenges with UNICEF’s internal procedures (administration, finance and procurement). The evaluation of UNICEF’s education in emergencies programming noted that the extent to which the evaluation could assess efficiency was limited because the program had failed to produce significant outputs and/or because of a lack of data. However, it did find that an efficient use of resources was reflected in the ability to leverage additional resources for the humanitarian response. On the other hand, there were delays in disbursement of funds in the first two years, which led to the reallocation of unspent funds from slow-performing country programs to those that were performing
better.\textsuperscript{34}

Similarly, the recent real-time assessment of interventions in the Sahel did not include information on cost efficiency but did note that there were delays in both the deployment of staff and funding. The delays in funding are attributed in the evaluation to the lack of a clear strategy for integrating other sectors (for example, WASH, education, child protection) into the nutrition response in the Sahel.\textsuperscript{35} The evaluation of the DFID-UNICEF program to strengthen UNICEF’s humanitarian capacity reflected positives findings with respect to the efficiency of the joint program. However, it also noted that “Attempts have been made from Headquarters to streamline and clarify finance and administrative systems for emergencies. However, at the operational level, these are yet to make any significant difference and delays in funds release, Programme Cooperation Agreements and supplies remain common, except in a few countries.”\textsuperscript{36} UNICEF staff report that these issues are currently being addressed through revisions to standard operating procedures for the different levels of emergencies. A number of other recent changes were outlined in the 2012 \textit{UNICEF Humanitarian Action for Children}.\textsuperscript{37}

3.6 Monitoring and Evaluation

3.6.1 Coverage

Coverage for two of the four sub-criteria for monitoring and evaluation (M&E) is strong – Sub-criterion 6.2 “Systems and processes for results monitoring and reporting are effective” (59 evaluations) and Sub-criterion 6.4 “Evaluation used to improve effectiveness” (61 evaluations) (Figure 17). Coverage for Sub-criterion 6.1 “Systems and processes for evaluation are effective” is moderate. However, coverage for Sub-criterion 6.3 “RBM systems are effective” is weak and, as such, the findings will not be included in this report. In assessing this sub-criterion, the reviewers were looking for evidence of a focus on results and the integration of results monitoring into management decision-making. This emphasis on a quite literal interpretation of the term RBM was developed so that the analysts could readily distinguish between findings on results monitoring and those relating to a more fully developed system for managing by results. This distinction


\textsuperscript{37} 2012 UNICEF Humanitarian Action for Children, p. 12 - 14
may, in turn, have resulted in a finding of “not addressed” for some evaluations, even if some elements of an RBM system were addressed in the evaluation report.

This was not found in sufficient evaluations, in spite of the emphasis in UNICEF on results-based programming.

Figure 17: Number of Evaluations Addressing Sub-criteria for Using Evaluation and Monitoring

![Diagram showing evaluation results]

3.6.2 Key Findings

The results for two sub-criteria are somewhat positive (Figure 18). The finding for Sub-criteria 6.1 “Systems and processes for evaluation are effective” and Sub-criterion 6.4 “Evaluation used to improve effectiveness” indicate that more than half the evaluations were rated as satisfactory or better, 64% and 57%, respectively. However, the findings for Sub-criterion 6.2 “Systems and processes for results monitoring and reporting are effective” are considerably less positive. Only 39% of the evaluations reflected a rating of satisfactory or better for this sub-criterion. As will be seen below, this is one of the factors detracting from the use of M&E.

As noted in Section 1.6, since 2011, UNICEF has made a significant effort to strengthen results definition, monitoring and reporting through the development and implementation of MoRES. The effect of this system on the strength of UNICEF monitoring will be clearer on completion of the planned formative evaluation of MoRES in 2013.
Sub-criteria 6.1: Effective Systems and Processes for Evaluation and 6.2: Effective Systems for Monitoring and Reporting

The findings for Sub-criterion 6.2 “Systems and processes for results monitoring and reporting are effective” tended to be somewhat more positive for programming in the MTSP focus area of Young Child Survival and Development. This is perhaps explained by the fact that many of these activities tend to be more operational and are long-standing UNICEF activities, which are easier to monitor than other areas. On the other hand, the findings for Sub-criterion 6.1 “Systems and processes for evaluation are effective” for Child Protection tend to be less positive than for other focus areas. This is possibly a reflection of the challenges of M&E in an area that is less operational.

The evaluations identify a range of UNICEF’s M&E tools including baseline studies, mid-term or formative evaluations, regular evaluations and other evaluative-type studies, such as costing studies and participation in lessons learned exercises and joint evaluations. While some evaluations noted the positive achievements with respect to M&E (for example, tools for M&E were well-established and regularly implemented and good quality reports are being produced) (Box 18), others reported weaknesses, including:

- Gaps in the M&E systems, including gaps in coverage, systems not fully operational, evaluations planned but not conducted and weaknesses in data collection and feedback from government and other partner organizations; and,

- Information from the M&E systems was of poor quality. This was somewhat
more likely to be identified in evaluations of programming in least-developed countries.

**Box 18: Evaluation and Implementation of Lessons Learned**

At the end of 2005, UNICEF conducted a major evaluation of the emergency response and initial phase (first six months after the tsunami) for Sri Lanka, Indonesia and Maldives. UNICEF also participated actively in the Tsunami Evaluation Coalition (TEC), which produced a series of evaluations and reports covering thematic topics, including: coordination; needs assessment; the impact of the international response on local and national capacities; links between relief, rehabilitation and development; and the funding response. In addition to this, there were regional consultations and ‘lessons learned’ exercises that captured some of the key findings. The recommendations and lessons from these evaluations have influenced adjustments in programme design and management, as well as the formulation of UNICEF’s emergency/early recovery response policies and capacities. More recently (end-2008), a follow-up Linking Relief Rehabilitation and Development (LRRD) study was undertaken by the Tsunami Evaluation Coalition.


Similar findings in:


Specific examples from the evaluations include:

- A series of evaluations were conducted of UNICEF’s malaria programming between 2003 and 2006 (Togo);
- Development of monitoring tools and teaching participants how to use field diaries, self-assessments and self-training for interventions to support the psychosocial development of children (Colombia);
- Development of an inexpensive and easy-to-use innovative inventory system for WASH programming that was successfully pilot tested (Ethiopia); and,
- Documentation of project processes from planning to planning, implementation, monitoring and evaluation that created a rich source of information to draw lessons learned and best practices.

A few evaluations noted specifically that these weaknesses meant that UNICEF could not measure the impact of its programming and that there was a tendency to focus on short-term accomplishments. This was noted in the global formative evaluation of the United Nations Girls’ Education Initiative: “Assessing the partnership’s impact in its outcome areas is challenging, in part because of the relatively short time frame for many of the initiatives, but also because of the absence of systems to track the impact, or the absence..."
of clear targets for what the various initiatives are meant to accomplish."

### Sub-criterion 6.4: Use of Evaluations to Improve Effectiveness

The results for Sub-criterion 6.4 reflect the extent to which UNICEF is preparing management responses for its evaluations and whether these include adequate responses to the recommendations, including an action plan and clear responsibility for its implementation. In 2008, the Executive Director was required by the Executive Board to ensure the preparation and availability of management responses for all evaluation reports and this requirement has been implemented gradually over the past few years. As a result, evaluations reflecting a satisfactory or better rating for this sub-criterion are much more likely to be found for evaluations published in 2010 and 2011, than in 2009. The data also suggest that adequate management responses are more likely to be found for evaluations of programming in MICs.

#### 3.6.3 Contributing Factors

The reviews identified very few factors that contributed positively to monitoring and evaluation. There were one or two references to the fact that the evaluations were part of a broader process of evaluation – either as part of UNICEF’s evaluation of country programs, or as part of a donor-led global evaluation. In one case it was noted that self-evaluations were conducted in anticipation of an external evaluation. A few evaluations noted that the need for data for program management contributed to the development of a monitoring system.

However, a range of factors were identified that detract from UNICEF’s ability to monitor and evaluate its programming adequately. Some evaluation reflected that:

- The lack of a clear results framework that would allow for monitoring results. This was reflected in the lack of logframes or logic models of programming or a disconnect between existing frameworks and the actual programming (Box 19); and,
- M&E was hampered by inadequate or inappropriate indicators and inadequate baseline information (Box 20).

---


Box 19: Lack of focus on results

“Management for results remains a major challenge. There is … a growing commitment to monitoring girls’ education processes and outputs but with less attention being paid to outcomes and impact that can be directly attributed to UNGEI. Results continue to be advanced in terms of (i) commitment of stakeholders; (ii) processes such as community participation at the investment stage; and (iii) direct project outputs such as girls’ and boys’ access and retention rates. Less interest has been directed to outcomes such as (a) community participation at the operational/implementation stage; and (b) the effects of training on actual teaching practices and how girls experience school differently as a result of teachers’ training. Strong beliefs in UNGEI’s effectiveness are thus not always backed by strong objective evidence.”


Box 20: Challenges with baseline data and indicators

“Measuring the results of the interventions is a challenge because of the absence of baseline data and that the statement of goals and objectives of all the four projects could have been more results-oriented and formulated in “SMART” way.”

Evaluation of Adolescent and Youth Participation in UNICEF Cambodia, 2011, [no page number]

A few reflected that:

- The planned M&E system was unrealistic, noting specifically the lack of adequate financial resources to implement the system;

- The structures for M&E, including the lack of clearly assigned responsibility for M&E and the lack of financial resources and adequately trained staff for carrying out M&E activities; and,

- A few noted the lack of M&E tools (for example, operational manuals, standards for M&E systems, reporting templates and quality assurance processes); and

- Factors that were reported in only one or two evaluations include:
  - Issues with follow-up, including the lack of mechanisms for lessons learned;
  - Programming being too ambitious or supervisions sites too spread out, leading to difficulties for monitoring; and,
  - Reluctance of staff to report non-compliant performance.
4.0 Conclusions

1. UNICEF-supported programs are highly relevant to the needs of target group members and are supportive of the development plans and priorities of program countries (satisfactory or better in almost all evaluations). UNICEF has also had success in developing effective partnerships with government and NGOs, especially in responding to humanitarian situations – which has, in itself, contributed to program relevance. The relevance of UNICEF programming has also been supported by efforts to ensure alignment with key national development plans and policy documents. It has also benefited from careful use of research into country conditions and the specific needs of target group members.

2. UNICEF has largely been effective in achieving the objectives of the development and humanitarian programs it supports, with three-quarters of evaluations reporting findings of satisfactory or better. UNICEF-supported programs have also been effective in securing positive benefits for target group members and in supporting positive changes in national policies and programs. Where UNICEF programs have achieved success in objectives achievement it has often been supported first by high quality program design with clear objectives that were integrated into national programs. It was also supported by UNICEF’s active role in influencing national sector policies. On the other hand, when UNICEF-supported programs failed to meet their objectives, it was most often due to weaknesses in project design, often linked to unclear causal relationships and lack of a results orientation.

3. These weaknesses in project design were contributing factors to mixed results with respect to sustainability. UNICEF achieves fairly positive ratings (more than two-thirds satisfactory or better) for its contributions to strengthening the enabling environment for development. However, the results for the likely continuation of results or the development of institutional and community capacity are not as good. Less than half the evaluations were rated as satisfactory or better on these two sub-criteria. A key factor to explain these results is the failure to plan for sustainability and integrate sustainability into program designs.

4. UNICEF’s performance with respect to gender equality is a serious concern. About one-third of evaluations did not address gender equality and, for those that did address it, the results were weak. As a result, only one-third of the evaluations were able to demonstrate that UNICEF-supported programs effectively address gender equality. A major factor contributing to these poor results was the lack of specific objectives regarding gender equality or the lack of a gender perspective in program design and implementation. Given the identification of gender equality as a
foundation strategy, it is surprising that gender is adequately addressed in only two-thirds of the reviewed evaluations.

5. There was insufficient coverage of environmental sustainability to warrant presentation of the results. It appears that environmental sustainability or the impact of UNICEF-supported programs on their environment is not addressed in most evaluations, although some evaluations of humanitarian action and programs in water, sanitation and hygiene did address the issue. Given increasing emphasis on programs to mitigate the effects of climate change in some UNICEF country programs, coverage of environmental sustainability may be expected to improve in future evaluation.

6. The results for efficiency of programming are mixed and the interpretation of the results is difficult because the sub-criteria associated with efficiency are not covered systematically in all evaluations. It is likely that factors related specifically to timeliness and implementation systems are only addressed in evaluations if they are problematic and can help to explain weakness in objectives achievement. However, the sub-criterion related to cost efficiency was more likely to be covered in the evaluations and showed somewhat positive results – almost two-thirds received a rating of satisfactory or better. However, this reflects ways in which programs have tried to reduce unit costs and increase the efficiency of resource use, rather than an analysis of overall program costs, as these costs were not identified in nearly half the evaluations. The factors that enhance cost efficiency include the establishment of effective monitoring systems to track costs and improve efficiency through lower delivery costs or unit prices and efforts to identify low-cost approaches to programming. However, where the results are not so positive, the detracting factor is most often the lack of appropriate cost data.

7. The evaluations reflected somewhat positive findings with respect to the use of evaluation at UNICEF, but less so for monitoring and reporting. Nearly two-thirds of the evaluations were rated as satisfactory or better in relation to UNICEF’s systems and processes for evaluation; but only two-fifths were given the same rating for UNICEF’s monitoring and reporting on effectiveness. The use of evaluations is supported by an increasing tendency to prepare management responses for evaluations that include action plans for the implementation of evaluation recommendations. There were insufficient ratings for the use of results-based management systems to report on the results. The lack of clear results frameworks and appropriate indicators and baseline information were factors that detracted from UNICEF’s effective use of monitoring and evaluation systems. The current development and implementation of the MoRES represents a significant effort to
address this issue.

8. While the sample of evaluations reviewed reflects reasonably the population of
UNICEF evaluations between 2009 and 2011, a review of the profile, by country and
by type of programming, of the evaluations conducted suggests that UNICEF does
not have adequate coverage of the programming in the countries that received the
largest amounts of both development and humanitarian funding. There is noticeably
limited coverage of UNICEF’s humanitarian action.

9. UNICEF is investing considerable time and effort into the development of quite new
systems for results monitoring which are being implemented at all levels of the
organization and in using the results of global evaluations for strategic planning.
These initiatives hold out the promise of strengthened results reporting which could
negate the need for development effectiveness reviews of this type in the future. That
could depend on how well the evaluation function is able to be incorporated into the
system as a means of verifying UNICEF’s contribution and testing the validity of
theories of change.

As noted in Section 3.0, UNICEF has undertaken a number of initiatives in the period
following publication of the evaluations reviewed, in order to address some of the issues
reported here. Examples of the main initiatives include:

- The development and ongoing implementation of the MoRES system for
defining, monitoring and reporting on program results (see Section 1.6) in concert
with the development of the Virtual Information Reporting System (VISION);

- The continued development by the Evaluation Office of the Global Evaluation
Report Oversight System (GEROS), coupled with efforts to improve evaluation
coverage;

- Changes to operating procedures to reduce procedural bottlenecks contributing to
programmatic delays, especially for humanitarian actions;

- Development of the Strategic Priority Action Plan for Gender Equality: 2010-
2012; and,

- Efforts to use evaluation and research findings to inform and strengthen the
development of the new strategic plan and thereby to improved effectiveness in
development programming and humanitarian action.

While the review did not carry out an assessment of the expected results of these
initiatives, it was able to verify from the documents reviewed and the interviews carried
out that they represent a significant effort to respond to the issues identified. Their
effectiveness will doubtless be the subject of future evaluations.
Annexes
Annex 1: Effectiveness Criteria

<table>
<thead>
<tr>
<th><strong>Relevance</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Programs are suited to the needs of target group members</td>
</tr>
<tr>
<td>1.2 Programs are aligned with national humanitarian and development goals</td>
</tr>
<tr>
<td>1.3 Effective partnerships with government, civil society and humanitarian and development partners</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Achievement of Humanitarian and Development Objectives and Expected Results</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Programs and projects achieve stated objectives</td>
</tr>
<tr>
<td>2.2 Positive benefits for target group members</td>
</tr>
<tr>
<td>2.3 Substantial numbers of beneficiaries/contribution to national humanitarian and development goals</td>
</tr>
<tr>
<td>2.4 Significant changes in national development policies/programs</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Cross Cutting Themes: Inclusive Humanitarian and Development Which can be Sustained (Gender Equality and Environmental Sustainability)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Programs effectively address gender equality</td>
</tr>
<tr>
<td>3.2 Changes are environmentally sustainable</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Sustainability</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Program benefits are likely to continue</td>
</tr>
<tr>
<td>4.2 Programs support institutional and community capacity</td>
</tr>
<tr>
<td>4.3 Programs strengthen enabling environment for humanitarian and development</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Efficiency</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Program activities are cost efficient</td>
</tr>
<tr>
<td>5.2 Programs are implemented/objectives achieved on time</td>
</tr>
<tr>
<td>5.3 Systems for program implementation are efficient</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Using Evaluation and Monitoring to Improve Humanitarian and Development Effectiveness</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1 Systems and processes for evaluation are effective</td>
</tr>
<tr>
<td>6.2 Systems and processes for monitoring are effective</td>
</tr>
<tr>
<td>6.3 Results based management systems are effective</td>
</tr>
<tr>
<td>6.4 Evaluation results used to improve humanitarian and development effectiveness</td>
</tr>
</tbody>
</table>
Annex 2: Evaluation Sample

7. 2009 Colombia: Educación en el Riesgo de Minas (ERM) y Asistencia a Víctimas en los departamentos de Cauca, Chocó, Nariño y la región de La Mojana /Sur de Bolívar
8. Evaluación Del Programa "Escuela Busca al Niño"
10. L’Evaluation des résultats des 12 Associations de Services Financiers (ASF) pour le développement Communautaire dans la zone de Kissidougou
11. Evaluation of the actions to attend to the water sanitation and hygiene sector among populations affected by Hurricane Felix
12. Evaluation de la couverture de la campagne nationale de distribution des moustiquaires imprégnées en 2008 et de l’impact des interventions de lutte contre le paludisme au Togo/Evaluation of the impact of anti malaria interventions, including 2008 national campaign to distribute ITNs
13. Evaluation of the UNICEF-DIPECHO programme “Supporting Disaster Risk Reduction Amongst Vulnerable Communities and Institutions in Central Asia and South Caucasus
14. Evaluation of the impact of educational innovations (IECD centers, satellite schools, non formal basic education centers) on Burkina Faso’s educational development system
15. Monitoring & Evaluation of Child Health Days in Madagascar
16. Mid-term Evaluation of project of Basic Education in Eastern DRC (BEED)
17. Recovery Action and Rehabilitation Project (RARP) Evaluation
18. Evaluation of the National Plan for the Eradication of Child Labour 2000-2010

20. 2010 Ethiopia: Mid-Term Evaluation of EU/UNICEF Supported WASH Programme


22. Evaluation, consolidation and strengthening of the “I'm a Person, Too” programme for optimum psychosocial development of children under six in the Department of Huila, Colombia - Evaluación del PROGRAMA “TAMBIÉN SOY PERSONA”

23. Evaluation of the YCSD programme

24. Evaluation of community-led total sanitation (CLTS)

25. Evaluation of the community-based rehabilitation project for disabled children in 4 municipalities of Oaxaca


30. 2009 Cambodia: Evaluation of Community-Led Total Sanitation

31. Community Radio Listening Groups Project

32. Evaluation of the EthioInfo Utilization in Ethiopia

33. Evaluation of PHAST tool for the promotion hygiene Sanitation in the GOK/UNICEF Programme of cooperation

34. 2010 Ghana: Review of Second Performance Monitoring of the IWASH Project

35. Evaluation of Integrated Management of Childhood Illnesses Initiative in the Republic of Moldova Years 2000-2010

36. Evaluation of Adolescent and Youth Participation in UNICEF Cambodia

37. Evaluation of Child and Youth Participation Initiatives in UNICEF Mozambique


39. Emergency Water Supply to un-served/underserved/ Vulnerable Areas in Baghdad and the IDPs

68
41. 2009 Uganda: Final Review of UNICEF-supported Programmes for Children Affected by Conflict in Kitgum, Northern Uganda
42. Supporting Sustainable water management and governance for the poor in drought and flood prone areas in Kenya
43. 2010 WCARO Regional: Roll Out Evaluation of Community Led Total Sanitation in West & Central
44. Evaluation of Gender Sensitisation and People Friendly Police Initiative in Karnataka
45. 2010 Iraq: Evaluation of Water Quality Control and Surveillance in Iraq
47. 2010 Sudan: Go To School Evaluation
49. Evaluation of Social Work Coaching
50. Evaluation of Sustained Outreach Services (SOS) for immunization/ Vitamin A (Indonesia)
53. 2010 Madagascar: Evaluation on girl-to-girl mentorship strategy
55. Evaluation of Netherlands-UNICEF Water Initiative (NUWI)
56. Real-Time Evaluation of UNICEF’s Response to the Sa’ada Conflict in Northern Yemen
58. Formative Evaluation of the United Nations Girls Education Initiative (UNGEI)
59. 2011 Mozambique: Impact evaluation of drinking water supply and sanitation interventions in rural Mozambique: More than Water
60. Assessing of WASH package interventions in 5 counties of Liberia
Annex 3: Methodology

This annex provides a more detailed explanation of the population identification and sampling methodology used for the review of UNICEF’s development and humanitarian effectiveness and includes a comparison of the sample to the population of evaluations and, on the issue of UNICEF funding, to global funding for the 2009 – 2011 period.

1. UNICEF Evaluation Population

UNICEF evaluations were identified from two sources:

- Office of Evaluation website (www.unicef.org/evaldatabase/); and,
- Three oversight reports from the assessment of evaluations in UNICEF’s Global Evaluation Reports Oversight System (GEROS), covering evaluations conducted from 2009 to the end of 2011.

This included evaluations conducted or commissioned by the central Office of Evaluation, as well as evaluations conducted at the decentralized level by UNICEF Country or Regional Offices.

These sources identified a population of 341 evaluations. It was decided, in consultation with the UNICEF EO that the review should focus on evaluations conducted since 2008. UNICEF had implemented a new evaluation policy in 2008 and evaluations conducted after that year would be likely to reflect the new policy. As a result, the scope of the review would include only evaluations conducted in 2009, 2010 and 2011.\(^\text{40}\) It was also agreed that the review should not include those evaluations deemed, in the GEROS quality rating, to be of poor quality and “not confident to act.” Finally, evaluations that focused on global programming were dropped from the population of evaluations from which to draw the sample for the qualitative review. These global evaluations were not included in the sample, but were rather subject to a qualitative review by senior team members. In addition, some evaluations in the population were duplicates, were not deemed to be evaluations or did not focus on UNICEF programming and were dropped from the population. This resulted in a population of 197 from which to draw the sample (Figure 19).

---

\(^{40}\) As the sample was drawn in 2012, not all evaluations conducted in that year were available and, as a result, 2012 evaluations were not included in the scope.
2. **Evaluation Sample**

A sample of 70 evaluations was drawn from the population, with the intent that, after the quality review, there would be 60 evaluations available for review. Initially the sample was random, stratified by MTSP theme area. Then the sample was adjusted manually to ensure adequate coverage on two other dimensions: region and type of country (low- or middle-income countries). The resulting sample was no longer random, but rather a purposive sample that was illustrative of UNICEF’s programming across a number of dimensions – year, commissioning office, region, type of country, country and MTSP focus area.

3. **Quality Review**

The first task of each reviewer was to review the quality of the evaluation, using a standard review grid (Figure 20). The grid reflects the criteria being rated and how the maximum number of points is allocated for each criterion.

The purpose of this review was two-fold:

- To ensure that the evaluations being used to provide information on UNICEF programming were of sufficiently overall high quality to be credible evaluations. This resulted in an overall quality score that had a maximum of 40 points. An evaluation had to score a minimum of 25 points to be included in the review; and,

- To ensure that, even if the evaluation was generally of high quality, the ratings for Criteria G, H and I were sufficiently high to ensure the evaluation would provide solid information specifically with respect to measuring effectiveness. If an

---

41 Some evaluations were dropped for more than one reason, so the total number of evaluations dropped, by reason, exceeds the number of evaluations dropped.
evaluation did not include sufficient lines of evidence (Criteria G), was not based on an adequate evaluation design (Criteria H) or included evaluation findings and conclusions that were not relevant and evidence based (Criteria I), then it was deemed that it was not likely to provide sufficient evidence with respect to development effectiveness. A total of 13 points was available for these three criteria and an evaluation had to receive a minimum of nine points in on these three criteria.

**Figure 20: Quality Review Grid**

<table>
<thead>
<tr>
<th>Points for criteria scored</th>
<th>Maximum Points</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Purpose of the evaluation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• why the evaluation was done (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• what triggered the evaluation (including timing in the project/program cycle) (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• how evaluation is to be used (1)</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>B Evaluation objectives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• evaluation objectives are clearly stated (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• objectives logically flow from purpose (1)</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>C Organization of the evaluation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• logical structure to the organization of the evaluation (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• evaluation is well written (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• clear distinction between evidence, findings, conclusions, and recommendations (1)</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>D Subject evaluated is clearly described</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluation describes:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• the activity/program being evaluated (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• the program’s expected achievements (for humanitarian: identification of needs of target population and agency specific role in addressing needs) (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• how the program addresses the development problem (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• the implementation modalities used (1)</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>E Scope of the evaluation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluation defines the boundaries of the evaluation in terms of:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• time period covered (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• implementation phase under review (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• geographic area (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• dimensions of stakeholder involvement being examined (1)</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>F Evaluation criteria</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluation criteria include:</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Points for criteria scored</td>
<td>Maximum Points</td>
<td>Score</td>
</tr>
<tr>
<td>---------------------------</td>
<td>----------------</td>
<td>-------</td>
</tr>
<tr>
<td>• the achievement of development objectives and expected results (including impacts) (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• cross-cutting issues: inclusive development which is gender sensitive and environmentally sustainable (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• the sustainability of benefits and positive results achieved (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• the relevance of MO activities and supported projects and programs (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• the efficiency of MO operations in support of projects and programs (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Multiple lines of evidence</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• one point (1) for each line of evidence used (case studies, surveys, site visits, interviews, etc.), up to a maximum of four points (4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Evaluation design</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elements of a good evaluation design include:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• an explicit theory of how objectives and results were to be achieved (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• specification of the level of results achieved (output, outcome, impact) (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• baseline data (quantitative or qualitative) on conditions prior to program implementation (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• comparison of conditions after program delivery to those before (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• a qualitative or quantitative comparison of conditions among program participants and a control group (not applicable to humanitarian evaluation, assign 1 point) (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Evaluation findings and conclusions are relevant and evidence based</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluation report includes:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• evaluation findings relevant to the assessment criteria (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• findings that are supported by the chosen methodologies (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• a clear logical link between the evidence and the finding (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• conclusions which are linked to the evaluation findings as reported (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Evaluation limitations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• statement of the limitations of the methodology (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• impact of limitations on evaluation (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• remedies of limitations (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Evaluation Recommendations</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*OECD DAC Network on Development Evaluation*
Review of UNICEF’s Development Effectiveness

<table>
<thead>
<tr>
<th>Points for criteria scored</th>
<th>Maximum Points</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>• evaluation contains recommendations that flow from findings and conclusions (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• recommendations are directed to one or more organization (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• recommendations are aimed at improving development effectiveness (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total (required to have a minimum of 25 points)</strong></td>
<td>40</td>
<td></td>
</tr>
<tr>
<td><strong>Total for Criteria G, H and I (required to have minimum of 9 points)</strong></td>
<td>13</td>
<td></td>
</tr>
</tbody>
</table>

Any evaluation that failed to meet either one of these conditions was rejected from the sample. The initial sample of evaluations included 70 evaluations. Of these four were rejected because:

- On detailed review, one was found to not address UNICEF programming and one was found to not address programming effectiveness; and,
- Two evaluations covered the same programming as other evaluations already included in the sample.

Of the remaining 66 evaluations, all achieved sufficient points in the overall quality score, but four did not receive sufficient points for the specific criteria related to measuring effectiveness and were rejected from the sample (Figure 21).

**Figure 21: Results of Quality Review Scoring**

<table>
<thead>
<tr>
<th>Overall evaluation quality scores</th>
<th>Development effectiveness key criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max points = 40 (Min required = 25)</td>
<td>Max points = 13 (Min required = 9)</td>
</tr>
<tr>
<td>No. of evaluations</td>
<td>% of evaluations</td>
</tr>
<tr>
<td>-------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>36 - 40</td>
<td>15</td>
</tr>
<tr>
<td>31 - 35</td>
<td>36</td>
</tr>
<tr>
<td>26 - 30</td>
<td>15</td>
</tr>
<tr>
<td>21 - 25</td>
<td></td>
</tr>
<tr>
<td>16 - 20</td>
<td></td>
</tr>
<tr>
<td>11 - 15</td>
<td></td>
</tr>
<tr>
<td>6 - 10</td>
<td></td>
</tr>
<tr>
<td>0 - 5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>66</td>
</tr>
</tbody>
</table>
4. Comparison of Evaluation Population and Sample

This section provides a profile of the evaluation population and sample and comments on the extent to which the sample is illustrative of the population of UNICEF evaluations and UNICEF programming for the period 2009 – 2011.

**Evaluation and Programming Years**

The sample was spread across the three years 2009 – 2011 (Figure 22). The split by year parallels that of the population of evaluations.

**Figure 22: Profile of Population and Sample of Evaluations, by Year**

<table>
<thead>
<tr>
<th>Year</th>
<th>Population</th>
<th></th>
<th>Sample</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
<td>Number</td>
<td>Percent</td>
</tr>
<tr>
<td>2009</td>
<td>75</td>
<td>38%</td>
<td>22</td>
<td>35%</td>
</tr>
<tr>
<td>2010</td>
<td>61</td>
<td>31%</td>
<td>23</td>
<td>37%</td>
</tr>
<tr>
<td>2011</td>
<td>61</td>
<td>31%</td>
<td>17</td>
<td>27%</td>
</tr>
<tr>
<td>Total</td>
<td>197</td>
<td>100%</td>
<td>62</td>
<td>100%</td>
</tr>
</tbody>
</table>

However, the programming covered by these evaluations spanned the period 2000 to 2010. Just over one-third of the evaluations covered three-year programs and another 40% four- to five-year programs (Figure 23). However, a few (16%) covered programs that lasted between six and eleven years.

**Figure 23: Profile of Sample of Evaluations, by Range of Years Covered**

<table>
<thead>
<tr>
<th>Span of Years Covered by Evaluations</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - 3 years</td>
<td>23</td>
<td>37.1%</td>
</tr>
<tr>
<td>4 - 5 years</td>
<td>25</td>
<td>40.3%</td>
</tr>
<tr>
<td>6 - 8 years</td>
<td>6</td>
<td>9.7%</td>
</tr>
<tr>
<td>9 - 11 years</td>
<td>5</td>
<td>8.1%</td>
</tr>
<tr>
<td>Not available</td>
<td>3</td>
<td>4.8%</td>
</tr>
<tr>
<td>Total</td>
<td>62</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

This represents a limitation of evaluation meta-syntheses. Evaluations are, by definition, retrospective and a meta-synthesis is even more retrospective, as it is based on a body of evaluations conducted over a much earlier period of time. UNICEF’s policies, strategies and approaches to programming have changed over these years, but the changes will not be reflected in all the evaluations. For this reason, the findings may be somewhat dated. To the extent possible, the review addresses this through observations gleaned from

---

42 Information is not available about the population of evaluations to permit a comparison.
recent interviews with UNICEF staff and a review of UNICEF documents.

**Commissioning Office**

One-fifth of the evaluations included in the sample were commissioned by the Central Evaluation Office (Figure 24). The reviewers were unable to identify the commissioning office in one-third of the evaluations (not available). Nearly half the evaluations (46%) were commissioned by UNICEF country or regional offices.

**Figure 24: Profile of Sample of Evaluations, by Commissioning Office**

<table>
<thead>
<tr>
<th>Evaluation Commission Office</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central Evaluation Office</td>
<td>12</td>
<td>19.4%</td>
</tr>
<tr>
<td>Country Office</td>
<td>27</td>
<td>43.5%</td>
</tr>
<tr>
<td>Regional Office</td>
<td>2</td>
<td>3.2%</td>
</tr>
<tr>
<td>Not available</td>
<td>21</td>
<td>33.9%</td>
</tr>
<tr>
<td>Total</td>
<td>62</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

**Regions and Type of Country**

The sample was spread across all regions and types of countries in which UNICEF is programming (Figures 25 and 26).

**Figure 25: Profile of Population and Sample of Evaluations, by Region**

<table>
<thead>
<tr>
<th>Region</th>
<th>Population</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
</tr>
<tr>
<td>Middle East and North Africa</td>
<td>18</td>
<td>9%</td>
</tr>
<tr>
<td>West and Central Africa</td>
<td>24</td>
<td>12%</td>
</tr>
<tr>
<td>Eastern and Southern Africa</td>
<td>44</td>
<td>22%</td>
</tr>
<tr>
<td>Latin America and Caribbean</td>
<td>22</td>
<td>11%</td>
</tr>
<tr>
<td>CEE/CIS</td>
<td>28</td>
<td>14%</td>
</tr>
<tr>
<td>South Asia</td>
<td>18</td>
<td>9%</td>
</tr>
<tr>
<td>East Asia and Pacific</td>
<td>29</td>
<td>15%</td>
</tr>
<tr>
<td>Multi-region</td>
<td>14</td>
<td>7%</td>
</tr>
<tr>
<td>Total</td>
<td>197</td>
<td>100%</td>
</tr>
</tbody>
</table>


**Figure 26: Profile of Population and Sample of Evaluations, by Country Classification**

<table>
<thead>
<tr>
<th>Country Classification</th>
<th>Population</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
</tr>
<tr>
<td>Least-developed countries</td>
<td>76</td>
<td>39%</td>
</tr>
<tr>
<td>Middle-income countries</td>
<td>90</td>
<td>46%</td>
</tr>
<tr>
<td>Regional/multi-country</td>
<td>12</td>
<td>6%</td>
</tr>
</tbody>
</table>
The sample included slightly more evaluations from West and Central Africa and slightly fewer evaluations from Central and Eastern Europe/Commonwealth of Independent States (CES/CIS) than would have been suggested by the population. As would be consistent with this, the sample included slightly more evaluations from least-developed countries that would have been suggested by the evaluation population.

**Countries**

This oversampling of least-developed countries in West and Central Africa occurred because the review had to oversample from the countries receiving the largest amount of UNICEF funding. Figure 27 reflects the countries that were in the top funded countries for development programming in 2007 – 2011. Nineteen evaluations (7% of all evaluations in the population) and nine evaluations in the sample (18%) come from the top funded countries of the five-year period 2007 - 2011. This suggests that UNICEF is not achieving strong evaluation coverage in the countries that receive the largest amounts of development funding. The team oversampled from the top funded countries to increase the coverage of countries that received the most UNICEF development funding.

**Figure 27: Profile of Population and Sample of Development Evaluations, by Funding and Country, by Funding for 2007 – 2011**

<table>
<thead>
<tr>
<th>Country</th>
<th>No. of years country included in top funded countries with development funding* (Maximum 5 year)</th>
<th>Evaluation identified (2009 – 2011)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Population (182 evaluations)</td>
</tr>
<tr>
<td>India</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Nigeria</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Ethiopia</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Afghanistan</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>

* The report identifies “Developing countries” and “Least developed countries”, which are included in the list of “Developing countries.” to identify the middle-income countries, the least-developed countries were removed from the list of “Developing countries” and these were deemed to be middle-income countries.

**OECD DAC Network on Development Evaluation**
There is a similar pattern with respect to humanitarian action. Figure 28 reflects the countries that were in the top funded countries for humanitarian action in 2007 – 2011. Seven evaluations in the population (47%) and three evaluations in the sample (23%) came from the countries that received the most humanitarian funding in the years 2007 – 2011. The coverage of the top funded countries is better for humanitarian action than development programming because the population of evaluations included Inter-Agency Steering Committee evaluations, which do not address UNICEF programming specifically. However, this suggests, as with development evaluations, that UNICEF is not achieving strong evaluation coverage of its programming in the countries that receive the largest amounts of humanitarian funding. The team oversampled from the top countries to increase the coverage of countries that received the most UNICEF humanitarian funding.

Figure 28: Profile of Population and Sample of Humanitarian Evaluations, by Funding and Country, by Funding for 2007 - 2011

There is a similar pattern with respect to humanitarian action. Figure 28 reflects the countries that were in the top funded countries for humanitarian action in 2007 – 2011. Seven evaluations in the population (47%) and three evaluations in the sample (23%) came from the countries that received the most humanitarian funding in the years 2007 – 2011. The coverage of the top funded countries is better for humanitarian action than development programming because the population of evaluations included Inter-Agency Steering Committee evaluations, which do not address UNICEF programming specifically. However, this suggests, as with development evaluations, that UNICEF is not achieving strong evaluation coverage of its programming in the countries that receive the largest amounts of humanitarian funding. The team oversampled from the top countries to increase the coverage of countries that received the most UNICEF humanitarian funding.

Figure 28: Profile of Population and Sample of Humanitarian Evaluations, by Funding and Country, by Funding for 2007 - 2011

There is a similar pattern with respect to humanitarian action. Figure 28 reflects the countries that were in the top funded countries for humanitarian action in 2007 – 2011. Seven evaluations in the population (47%) and three evaluations in the sample (23%) came from the countries that received the most humanitarian funding in the years 2007 – 2011. The coverage of the top funded countries is better for humanitarian action than development programming because the population of evaluations included Inter-Agency Steering Committee evaluations, which do not address UNICEF programming specifically. However, this suggests, as with development evaluations, that UNICEF is not achieving strong evaluation coverage of its programming in the countries that receive the largest amounts of humanitarian funding. The team oversampled from the top countries to increase the coverage of countries that received the most UNICEF humanitarian funding.

There is a similar pattern with respect to humanitarian action. Figure 28 reflects the countries that were in the top funded countries for humanitarian action in 2007 – 2011. Seven evaluations in the population (47%) and three evaluations in the sample (23%) came from the countries that received the most humanitarian funding in the years 2007 – 2011. The coverage of the top funded countries is better for humanitarian action than development programming because the population of evaluations included Inter-Agency Steering Committee evaluations, which do not address UNICEF programming specifically. However, this suggests, as with development evaluations, that UNICEF is not achieving strong evaluation coverage of its programming in the countries that receive the largest amounts of humanitarian funding. The team oversampled from the top countries to increase the coverage of countries that received the most UNICEF humanitarian funding.

There is a similar pattern with respect to humanitarian action. Figure 28 reflects the countries that were in the top funded countries for humanitarian action in 2007 – 2011. Seven evaluations in the population (47%) and three evaluations in the sample (23%) came from the countries that received the most humanitarian funding in the years 2007 – 2011. The coverage of the top funded countries is better for humanitarian action than development programming because the population of evaluations included Inter-Agency Steering Committee evaluations, which do not address UNICEF programming specifically. However, this suggests, as with development evaluations, that UNICEF is not achieving strong evaluation coverage of its programming in the countries that receive the largest amounts of humanitarian funding. The team oversampled from the top countries to increase the coverage of countries that received the most UNICEF humanitarian funding.

<table>
<thead>
<tr>
<th>Country</th>
<th>No. of years country included in top funded countries with development funding* (Maximum 5 year)</th>
<th>Evaluation identified (2009 – 2011)</th>
<th>Population (182 evaluations)</th>
<th>Sample (49 evaluations)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pakistan</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Mozambique</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Malawi</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Southern Sudan</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Somalia</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>19</td>
<td>9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As % of all evaluations in population or sample 6.6% 18.4%

* Regular resources and Other resources - Regular

There is a similar pattern with respect to humanitarian action. Figure 28 reflects the countries that were in the top funded countries for humanitarian action in 2007 – 2011. Seven evaluations in the population (47%) and three evaluations in the sample (23%) came from the countries that received the most humanitarian funding in the years 2007 – 2011. The coverage of the top funded countries is better for humanitarian action than development programming because the population of evaluations included Inter-Agency Steering Committee evaluations, which do not address UNICEF programming specifically. However, this suggests, as with development evaluations, that UNICEF is not achieving strong evaluation coverage of its programming in the countries that receive the largest amounts of humanitarian funding. The team oversampled from the top countries to increase the coverage of countries that received the most UNICEF humanitarian funding.

There is a similar pattern with respect to humanitarian action. Figure 28 reflects the countries that were in the top funded countries for humanitarian action in 2007 – 2011. Seven evaluations in the population (47%) and three evaluations in the sample (23%) came from the countries that received the most humanitarian funding in the years 2007 – 2011. The coverage of the top funded countries is better for humanitarian action than development programming because the population of evaluations included Inter-Agency Steering Committee evaluations, which do not address UNICEF programming specifically. However, this suggests, as with development evaluations, that UNICEF is not achieving strong evaluation coverage of its programming in the countries that receive the largest amounts of humanitarian funding. The team oversampled from the top countries to increase the coverage of countries that received the most UNICEF humanitarian funding.

<table>
<thead>
<tr>
<th>Country</th>
<th>No. of years country included in top funded countries with humanitarian funding* (Maximum 5 years)</th>
<th>Evaluation identified (2009 – 2011)</th>
<th>Population (15 evaluations)</th>
<th>Sample (13 evaluations)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Somalia</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pakistan**</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Northern Sudan</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ethiopia</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Southern Sudan</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

* A longer time period was taken for the financial data since evaluations conducted in 2009 – 2011 covered programming (and, hence, expenditures) in earlier years. 2007 was the earliest year for which the team had expenditure data.
Review of UNICEF’s Development Effectiveness

<table>
<thead>
<tr>
<th>Country</th>
<th>No. of years country included in top funded countries with humanitarian funding* (Maximum 5 years)</th>
<th>Evaluation identified (2009 – 2011)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Population (15 evaluations)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sample (13 evaluations)</td>
</tr>
<tr>
<td>Sri Lanka</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Indonesia</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>As % of all evaluations in population or sample</td>
<td>46.7%</td>
<td>23.1%</td>
</tr>
</tbody>
</table>

* Funding from Other resources – emergency

** Evaluations from Pakistan were not included in the sample because they did not cover UNICEF programming

**MTSP Focus Areas**

The evaluations in the sample covered all priority areas of UNICEF programming, including all themes in the MTSP and humanitarian action (see Figure 29). However, not all types of UNICEF programming are covered equally by the evaluations. Only fifteen evaluations in the population (8%) and thirteen (19%) in the sample covered humanitarian action; whereas this programming accounted for just over one-quarter of UNICEF expenditures in 2008 – 2011. There is proportionally less coverage of humanitarian action by UNICEF’s evaluations that development programming. It should also be noted that five of the humanitarian evaluations included in the sample, although not for programming in the same country, covered programming for the same humanitarian crisis – the 2004 Indian Ocean tsunami. Only one of the hardest hit countries receiving funding in response to the tsunami is included in the list of top humanitarian funded countries for 2009 – 2011. This means that, although the evaluations were published in years included in the sample, the funding occurred several years earlier.

**Figure 29: Profile of Population and Sample of Evaluations, by MTSP Focus Area and Humanitarian Action**

<table>
<thead>
<tr>
<th>MTSP Focus Area/Humanitarian Action</th>
<th>% of Expenditures (2009 – 2011)*</th>
<th>Population</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Number</td>
<td>Percent</td>
</tr>
<tr>
<td>Development (Regular and Other Resources - Regular)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Young Child Survival and Development</td>
<td>33.40%</td>
<td>43</td>
<td>22%</td>
</tr>
<tr>
<td>2. Basic Education and Gender Equality</td>
<td>16.40%</td>
<td>65</td>
<td>33%</td>
</tr>
<tr>
<td>3. HIV/AIDS and Children</td>
<td>5.20%</td>
<td>11</td>
<td>6%</td>
</tr>
<tr>
<td>4. Child Protection from Violence, Exploitation and Abuse</td>
<td>7.40%</td>
<td>35</td>
<td>18%</td>
</tr>
</tbody>
</table>

OECD DAC Network on Development Evaluation
### MTSP Focus Area/Humanitarian Action

<table>
<thead>
<tr>
<th>MTSP Focus Area/Humanitarian Action</th>
<th>% of Expenditures (2009 – 2011)*</th>
<th>Population</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Number</td>
<td>Percent</td>
</tr>
<tr>
<td>5. Policy Advocacy and Partnership</td>
<td>9.80%</td>
<td>17</td>
<td>9%</td>
</tr>
<tr>
<td>Sub-total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Humanitarian (Other Resources -</td>
<td>26.60%</td>
<td>15</td>
<td>8%</td>
</tr>
<tr>
<td>Emergency)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100.00%</td>
<td>15</td>
<td>8%</td>
</tr>
<tr>
<td>Other (Country Program, Other, Blank)</td>
<td></td>
<td>11</td>
<td>6%</td>
</tr>
<tr>
<td>Total Evaluations</td>
<td></td>
<td>197</td>
<td>100%</td>
</tr>
</tbody>
</table>

*Source: “Expenditures by MTSP”, UNICEF, November 2012. It is noted that the programming and hence the funding covered by the evaluations in the sample occurred, for the most part, in years preceding these years.

Apart from the fact that the evaluations do not cover the countries receiving the largest amounts of funding, or the full range of humanitarian action, the evaluations covered in the sample are, in all other aspects, illustrative of UNICEF’s global programming.
## Annex 4: Guide for Review Team to Classify Evaluation Findings

### Common Development and Humanitarian Evaluation Assessment Criteria

<table>
<thead>
<tr>
<th>Criteria</th>
<th>(1) Highly Unsatisfactory</th>
<th>(2) Unsatisfactory</th>
<th>(3) Satisfactory</th>
<th>(4) Highly Satisfactory</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Relevance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.1 Multilateral organization supported programs and projects are suited to the needs and/or priorities of the target group</td>
<td>Evaluation finds that substantial elements of program or project activities and outputs were unsuited to the needs and priorities of the target group.</td>
<td>Evaluation finds that no systematic analysis of target group needs and priorities took place during the design phase of developmental or relief and rehabilitation programming or the evaluation report indicates some evident mismatch between program and project activities and outputs and the needs and priorities of the target group.</td>
<td>Evaluation finds that the multilateral organization supported activity, program or project is designed taking into account the needs of the target group as identified through a process of situation or problem analysis (including needs assessment for relief operations) and that the resulting activities are designed to meet the needs of the target group.</td>
<td>Evaluation finds methods used in program and project humanitarian and development (including needs assessment for relief operations) to identify target group needs and priorities (including consultations with target group members) and finds that the program and project takes those needs into account and is designed to meet those needs and priorities (whether or not it does so successfully).</td>
</tr>
<tr>
<td>1.2 Multilateral organization supported humanitarian and development projects and programs align with national humanitarian and development goals:</td>
<td>The evaluation reports that significant elements of multilateral organization supported humanitarian and development program and project activity run counter to national humanitarian and development priorities with a resulting loss of effectiveness.</td>
<td>The evaluation reports a significant portion (1/4 or more) of the multilateral organization supported humanitarian and development programs and projects subject to the evaluation are not aligned with national plans and priorities but there is no evidence that they run counter to those priorities.</td>
<td>Most multilateral organization supported humanitarian and development programs and projects are reported in the evaluation to be fully aligned with national plans and priorities as expressed in national poverty eradication strategy and sector strategic priorities.</td>
<td>All multilateral organization supported humanitarian and development projects and programs subject to the evaluation are reported in the evaluation to be fully aligned to national humanitarian and development goals as described in national and sector plans and priorities, especially including the national poverty eradication strategy and sector strategic priorities.</td>
</tr>
</tbody>
</table>
### Criteria (1) Highly Unsatisfactory (2) Unsatisfactory (3) Satisfactory (4) Highly Satisfactory

**1.3 Multilateral organization has developed an effective partnership with governments, bilateral and multilateral development and humanitarian organizations and non-governmental organizations for planning, coordination and implementation of support to development and/or emergency preparedness, humanitarian relief and rehabilitation efforts.**

Evaluation finds that the multilateral organization experiences significant divergence in priorities from those of its (government, non-governmental organization or donor) partners and lacks a strategy or plan which will credibly address the divergence and which should result in strengthened partnership over time.

Evaluation finds that multilateral organization has experienced significant difficulties in developing an effective relationship with partners and that there has been significant divergence in the priorities of the multilateral organization and its partners.

Evaluation finds that multilateral organization has improved the effectiveness of its partnership relationship with partners over time during the evaluation period and that this partnership was effective at the time of the evaluation or was demonstrably improved.

Evaluation finds that multilateral organization has consistently achieved a high level of partnership during the evaluation period.

<table>
<thead>
<tr>
<th>2. Achievement of Humanitarian and Development Objectives and Expected Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2.1 MO supported programs and projects achieve their stated development and/or humanitarian objectives and attain expected results.</strong></td>
</tr>
<tr>
<td>Less than half of stated output and outcome objectives have been achieved including one or more very important output and/or outcome level objectives.</td>
</tr>
</tbody>
</table>

Problems in the design or delivery of MO supported activities mean that expected positive benefits for target group members have not occurred or are unlikely to occur.

MO supported projects and programs result in no or very few positive changes experienced by target group members. These benefits may include the avoidance or reduction of negative effects of a sudden onset or protracted emergency.

MO supported projects and programs have resulted in widespread and significant positive changes experienced by target group members as measured using either quantitative or qualitative methods (possibly including comparison of impacts with non-program participants). These benefits may include...
<table>
<thead>
<tr>
<th>Criteria</th>
<th>(1) Highly Unsatisfactory</th>
<th>(2) Unsatisfactory</th>
<th>(3) Satisfactory</th>
<th>(4) Highly Satisfactory</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3 MO programs and projects made differences for a substantial number of beneficiaries and where appropriate contributed to national development goals.</td>
<td>MO supported projects and programs have not contributed to positive changes in the lives of beneficiaries as measured quantitatively or qualitatively.</td>
<td>MO supported projects and programs have contributed to positive changes in the lives of only a small number of beneficiaries (when compared to project or program targets and local or national goals if established).</td>
<td>MO supported projects and programs have contributed to positive changes in the lives of substantial numbers of beneficiaries as measured quantitatively or qualitatively. These may result from development, relief, or protracted relief and rehabilitation operations and may include the avoidance of negative effects of emergencies.</td>
<td>MO supported projects and programs have contributed to positive changes in the lives of substantial numbers of beneficiaries. Further, they have contributed to the achievement of specific national development goals or have contributed to meeting humanitarian relief objectives agreed to with the national government and/or national and international development and relief organizations.</td>
</tr>
<tr>
<td>2.4 MO activities contributed to significant changes in national development policies and programs (including for disaster preparedness, emergency response and rehabilitation) (policy impacts) and/or to needed system reforms.</td>
<td>National policies and programs in a given sector or area of development (including disaster preparedness, emergency response and rehabilitation) were deficient and required strengthening but MO activities have not addressed these deficiencies.</td>
<td>MO activities have not made a significant contribution to the development of national policies and programs in a given sector or area of development, disaster preparedness, emergency response or rehabilitation. (Policy changes in humanitarian situations may include allowing access to the affected populations).</td>
<td>MO activities have made a significant contribution to either re-orienting or sustaining effective national policies and programs in a given sector or area of development disaster preparedness, emergency response or rehabilitation.</td>
<td>MO activities have substantial contribution to either re-orienting or sustaining effective national policies and programs in a given sector or area of development disaster preparedness, emergency response or rehabilitation. Further, the supported policies and program implementation modalities have resulted in improved positive impacts for target group members.</td>
</tr>
</tbody>
</table>

3. Cross Cutting Themes: Inclusive Humanitarian Assistance and Development Which can be Sustained

<table>
<thead>
<tr>
<th>3.1 Extent MO supported activities effectively address</th>
<th>MO supported activities are unlikely to contribute to</th>
<th>MO supported activities either lack gender equality</th>
<th>MO supported programs and projects achieve a majority</th>
<th>MO supported programs and projects achieve all or nearly...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criteria</td>
<td>(1) Highly Unsatisfactory</td>
<td>(2) Unsatisfactory</td>
<td>(3) Satisfactory</td>
<td>(4) Highly Satisfactory</td>
</tr>
<tr>
<td>----------</td>
<td>--------------------------</td>
<td>-------------------</td>
<td>-----------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>the cross-cutting issue of gender equality.</td>
<td>gender equality or may in fact lead to increases in gender inequalities.</td>
<td>objectives or achieve less than half of their stated gender equality objectives. (Note: where a program or activity is clearly gender-focused (maternal health programming for example) achievement of more than half its stated objectives warrants a satisfactory rating).</td>
<td>(more than 50%) of their stated gender equality objectives.</td>
<td>all of their stated gender equality objectives.</td>
</tr>
<tr>
<td>3.2 Extent changes are environmentally sustainable.</td>
<td>MO supported programs and projects do not include planned activities or project design criteria intended to promote environmental sustainability. In addition changes resulting from MO supported programs and projects are not environmentally sustainable.</td>
<td>MO supported programs and projects do not include planned activities or project design criteria intended to promote environmental sustainability. There is, however, no direct indication that project or program results are not environmentally sustainable. OR MO supported programs and projects include planned activities or project design criteria intended to promote sustainability but these have not been successful.</td>
<td>MO supported programs and projects include some planned activities and project design criteria to ensure environmental sustainability. These activities are implemented successfully and the results are environmentally sustainable.</td>
<td>MO supported programs and projects are specifically designed to be environmentally sustainable and include substantial planned activities and project design criteria to ensure environmental sustainability. These plans are implemented successfully and the results are environmentally sustainable.</td>
</tr>
</tbody>
</table>

### 4. Sustainability

<p>| 4.1 Benefits continuing or likely to continue after project or program completion or there are effective measures to link the humanitarian relief operations, to rehabilitation, reconstructions and, eventually, to longer-term developmental results. | There is a very low probability that the program/project will result in continued intended benefits for the target group after completion. For humanitarian relief operations, the evaluation finds no strategic or operational | There is a low probability that the program/project will result in continued benefits for the target group after completion. For humanitarian relief operations, efforts to link the relief phase to rehabilitation, reconstruction and, eventually, to development are | Likely that the program or project will result in continued benefits for the target group after completion. For humanitarian relief operations, the strategic and operational measures to link relief to rehabilitation, reconstruction and, eventually, development | Highly likely that the program or project will result in continued benefits for the target group after completion. For humanitarian relief operations, the strategic and operational measures to link relief to rehabilitation, reconstruction and, eventually, development |</p>
<table>
<thead>
<tr>
<th>Criteria</th>
<th>(1) Highly Unsatisfactory</th>
<th>(2) Unsatisfactory</th>
<th>(3) Satisfactory</th>
<th>(4) Highly Satisfactory</th>
</tr>
</thead>
<tbody>
<tr>
<td>measures to link relief, to rehabilitation, reconstruction and, eventually, to development.</td>
<td>inadequate. (Note, in some circumstances such linkage may not be possible due to the context of the emergency. If this is stated in the evaluation, a rating of satisfactory can be given)</td>
<td>are credible.</td>
<td>eventually, development are credible. Further, they are likely to succeed in securing continuing benefits for target group members.</td>
<td></td>
</tr>
<tr>
<td>4.2 Extent MO supported projects and programs are reported as sustainable in terms of institutional and/or community capacity.</td>
<td>The design of MO supported programs and projects failed to address the need to strengthen institutional and/or community capacity as required. In the case of humanitarian operations, the design of programs and projects failed to take account of identified needs to strengthen local capacities for delivery of relief operations and/or for managing the transition to rehabilitation and/or development.</td>
<td>MO programs and projects may have failed to contribute to strengthening institutional and/or community capacity or, where appropriate, to strengthen local capacities for delivery of relief operations and/or for managing the transition to rehabilitation and/or development.</td>
<td>MO programs and projects have contributed to strengthening institutional and/or community capacity but with limited success.</td>
<td>Either MO programs and projects have contributed to significantly strengthen institutional and/or community capacity as required or institutional partners and communities already had the required capacity to sustain program results.</td>
</tr>
<tr>
<td>4.3 Extent MO development programming contributes to strengthening the enabling environment for development.</td>
<td>For development programs, there were important weaknesses in the enabling environment for development (the overall framework and process for national development planning; systems and processes for public consultation and for participation by civil society in development planning; governance structures and the rule of law; national and local mechanisms for accountability for public expenditures.</td>
<td>MO development activities and/or MO supported projects and programs have not made a notable contribution to changes in the enabling environment for development.</td>
<td>MO development activities and/or MO supported projects and programs have made a notable contribution to changes in the enabling environment for development including one or more of: the overall framework and process for national development planning; systems and processes for public consultation and for participation by civil society in development planning; governance structures and the rule of law; national and local mechanisms for accountability for public expenditures.</td>
<td>MO development activities and/or MO supported projects and programs have made a significant contribution to changes in the enabling environment for development including one or more of: the overall framework and process for national development planning; systems and processes for public consultation and for participation by civil society in development planning; governance structures and the rule of law; national and local mechanisms for accountability for public expenditures.</td>
</tr>
<tr>
<td>Criteria</td>
<td>(1) Highly Unsatisfactory</td>
<td>(2) Unsatisfactory</td>
<td>(3) Satisfactory</td>
<td>(4) Highly Satisfactory</td>
</tr>
<tr>
<td>----------</td>
<td>---------------------------</td>
<td>-------------------</td>
<td>----------------</td>
<td>------------------------</td>
</tr>
<tr>
<td></td>
<td>service delivery and quality; and necessary improvements to supporting structures such as capital and labour markets. Further, the MO activities and support provided to programs and projects failed to address the identified weakness successfully, further limiting program results.</td>
<td>rule of law; national and local mechanisms for accountability for public expenditures, service delivery and quality; and necessary improvements to supporting structures such as capital and labour markets.</td>
<td></td>
<td>rule of law; national and local mechanisms for accountability for public expenditures, service delivery and quality; and necessary improvements to supporting structures such as capital and labour markets.</td>
</tr>
<tr>
<td></td>
<td>MO supported programs and projects under evaluation (development, emergency preparedness, relief and rehabilitation) do not have credible, reliable information on the costs of activities and inputs and therefore the evaluation is not able to report on cost/resource efficiency. OR MO supported programs and projects under evaluation present mixed findings on the cost/resource efficiency of the inputs.</td>
<td>Level of program outputs achieved (development, emergency preparedness, relief and rehabilitation) when compared to the cost of program activities and inputs are appropriate even when the program design process did not directly consider alternative program delivery methods and their associated costs.</td>
<td></td>
<td>MO supported (development, emergency preparedness, relief and rehabilitation) programs and projects are designed to include activities and inputs that produce outputs in the most cost/resource efficient manner available at the time.</td>
</tr>
<tr>
<td>5. Efficiency</td>
<td>5.1 Program activities are evaluated as cost/resource efficient: Credible information indicating that MO supported programs and projects (development, emergency preparedness, relief and rehabilitation) are not cost/resource efficient.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.2 Evaluation indicates implementation and objectives achieved on time (given the context, in the case of humanitarian programming)</td>
<td>Less than half of stated output and outcome level objectives of MO supported programs and projects are achieved on time, there is no credible plan or legitimate explanation found by the evaluation which would suggest significant</td>
<td>Less than half of stated output and outcome level objectives of MO supported programs and projects are achieved on time but the program or project design has been adjusted to take account of difficulties encountered and</td>
<td>More than half of stated output and outcome level objectives of MO supported programs and projects are achieved on time and that this level is appropriate to the context faced by the program during implementation,</td>
<td>Nearly all stated output and outcome level objectives of MO supported programs and projects are achieved on time or, in the case of humanitarian programming, a legitimate explanation for delays in the achievement of some</td>
</tr>
</tbody>
</table>
### Criteria

<table>
<thead>
<tr>
<th>Criteria</th>
<th>(1) Highly Unsatisfactory</th>
<th>(2) Unsatisfactory</th>
<th>(3) Satisfactory</th>
<th>(4) Highly Satisfactory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improvement in on-time objectives achievement in the future.</td>
<td>Can be expected to improve the pace of objectives achievement in the future.</td>
<td>Particularly for humanitarian programming.</td>
<td>Outputs/outcomes is provided.</td>
<td></td>
</tr>
<tr>
<td>5.3 Evaluation indicates that MO systems and procedures for project/program implementation and follow up are efficient (including systems for engaging staff, procuring project inputs, disbursing payment, logistical arrangements etc.)</td>
<td>Serious deficiencies in agency systems and procedures for project/program implementation that result in significant delays in project start-up, implementation or completion and/or significant cost increases.</td>
<td>Some deficiencies in agency systems and procedures for project/program implementation but does not indicate that these have contributed to delays in achieving project/program objectives.</td>
<td>Agency systems and procedures for project implementation are reasonably efficient and have not resulted in significant delays or increased costs.</td>
<td>Efficiency of agency systems and procedures for project implementation represent an important organizational strength in the implementation of the program under evaluation.</td>
</tr>
</tbody>
</table>

### 6. Using Evaluation and Monitoring to Improve humanitarian and development Effectiveness

<table>
<thead>
<tr>
<th>6.1 Systems and processes for evaluation are effective.</th>
<th>Evaluation practices in use for programs and projects of this type (development, emergency preparedness, relief and rehabilitation) are seriously deficient.</th>
<th>No indication that programs and projects of this type (development, emergency preparedness, relief and rehabilitation) are subject to systematic and regular evaluations.</th>
<th>Program being evaluated is subject to systematic and regular evaluations or describes significant elements of such practice. No mention of policy and practice regarding similar programs and projects. This may include specialized evaluation methods and approaches to emergency preparedness, relief and rehabilitation programming.</th>
<th>Program being evaluated (along with similar programs and projects) is subject to systematic regular evaluations or describes significant elements of such practice.</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.2 Systems and processes for monitoring and reporting on program results are effective.</td>
<td>Absence of monitoring and reporting systems for the development and humanitarian programming. This would include the absence of adequate monitoring and reporting systems.</td>
<td>While monitoring and reporting systems for the development and humanitarian programming exist, they either do not report on a regular basis or they are not well-established and report regularly.</td>
<td>Monitoring and reporting systems for development and humanitarian programming as appropriate are well-established and report regularly.</td>
<td>Monitoring and reporting systems for the program are well-established and report regularly. The quality of regular reports is rated highly by the evaluation and results.</td>
</tr>
<tr>
<td>Criteria</td>
<td>(1) Highly Unsatisfactory</td>
<td>(2) Unsatisfactory</td>
<td>(3) Satisfactory</td>
<td>(4) Highly Satisfactory</td>
</tr>
<tr>
<td>----------</td>
<td>--------------------------</td>
<td>-------------------</td>
<td>-----------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>monitoring of outputs during the implementation of humanitarian programming.</td>
<td>inadequate in frequency, coverage or reliability.</td>
<td></td>
<td></td>
<td>are reportedly used in the management of the program.</td>
</tr>
<tr>
<td>6.3 Results Based Management (RBM) systems are effective</td>
<td>No evidence of the existence of an RBM system for the program and no system is being developed.</td>
<td>While an RBM system is in place, or being developed, it is unreliable and does not produce regular reports on program performance.</td>
<td>RBM system is in place and produces regular reports on program performance.</td>
<td>RBM system is in place for the program and there is evidence noted in the evaluation that the system is used to make changes in the program to improve effectiveness.</td>
</tr>
<tr>
<td>6.4 MO makes use of evaluation to improve development/humanitarian effectiveness</td>
<td>Report does not include a management response and does not have one appended to it or associated with it. There is no indication of how the evaluation results will be used. There is no indication that similar evaluations have been used to improve effectiveness in the past.</td>
<td>Report includes a management response (or has one attached or associated with it) but it does not indicate which recommendations have been accepted. OR There is some, non-specific indication that similar evaluations have been used to improve program effectiveness in the past.</td>
<td>Report includes a management response (or has one attached or associated with it) that indicates which recommendations have been accepted. OR There is a clear indication that similar evaluations in the past have been used to make clearly identified improvements in program effectiveness.</td>
<td>Report includes a management response (or has one attached or associated with it) describes a response to each major recommendation which is appropriate and likely to result in the organizational and programmatic changes needed to achieve their intent.</td>
</tr>
</tbody>
</table>
Annex 5: Global Evaluations and UNICEF Documents

1. **UNICEF Global Evaluations**

Global Evaluation of DevInfo, 2009
A Study of UNICEF Engagement in Global Programme Partnerships, 2009
Evaluation of UNICEF Multiple Indicator Cluster Surveys Round 3 (MICS3), 2009
UNICEF Child Friendly Schools Programming Evaluation, 2009
Progress Evaluation of the UNICEF Education in Emergencies and Post-crisis Transition Programme, 2010
Review of the Global Education Cluster Co-leadership Arrangement, 2010
Unite for Children, Unite Against AIDS Campaign Evaluation, 2010
Final Report: Evaluation of UNICEF’s Programme and Work in Relation to Adolescents and the Participation of Children and Young People, 2010
Evaluation of UNICEF's Early Childhood Development Programme with Focus on Government of Netherlands Funding (2008-2010), 2011
5-Year Evaluation of the Central Emergency Response Fund - Final Synthesis Report, OCHA

2. **Other Documents**

**UNICEF Strategy and Policy Documents**
2012 UNICEF Humanitarian Action for Children, UNICEF
The State Of The World’s Children 2012: Children in an Urban World, UNICEF
UNICEF Evaluation Policy, UNICEF, 5 December 2007

*OECD DAC Network on Development Evaluation*
**UNICEF Reports**
UNICEF Compendium of decisions adopted by the Executive Board in 2008, 1 October 2008


Annual report of the Executive Direction: progress and achievements against the medium-term strategic plan, 19 April 2011

Annual report of the Executive Director of UNICEF: progress and achievements against the medium-term strategic plan, UNICEF, 2012

Annual report on the evaluation function and major evaluations in UNICEF, 13 July 2010

Annual report on the evaluation function and major evaluations in UNICEF, 3 April 2012


**UNICEF Monitoring and Evaluation Documents**

“MoRES Monitoring Results for Equity Systems: Access and Quality in Early Learning”, UNICEF, presentation [undated]


**OTHER UNICEF Documents**


UNICEF and the next MTSP: key opportunities and challenges for programming, Programme Division, UNICEF, 2012

**GEROS Documents**


Documents from Other Organizations


Annex 6: Selected Results for Comparisons of Least-Developed and Middle-Income Countries and by Focus Area and Humanitarian Action

1. Comparison of Least-developed and Middle-income Countries

1.3 Effective partnerships

<table>
<thead>
<tr>
<th>Country Type</th>
<th>(4) Highly Satisfactory</th>
<th>(3) Satisfactory</th>
<th>(2) Unsatisfactory</th>
<th>(1) Highly Unsatisfactory</th>
<th>Total</th>
<th>Valid Cases</th>
<th>Mean Score</th>
<th>Total Cases</th>
<th>Satisfactory or Better</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDC</td>
<td>28.1%</td>
<td>53.1%</td>
<td>9.4%</td>
<td>9.4%</td>
<td>100.0%</td>
<td>32</td>
<td>3.00</td>
<td>34</td>
<td>81%</td>
</tr>
<tr>
<td>MIC</td>
<td>27.3%</td>
<td>68.2%</td>
<td>4.5%</td>
<td>0%</td>
<td>100.0%</td>
<td>22</td>
<td>3.23</td>
<td>26</td>
<td>95%</td>
</tr>
</tbody>
</table>

2.1 Programs achieve their state objectives

<table>
<thead>
<tr>
<th>Country Type</th>
<th>(4) Highly Satisfactory</th>
<th>(3) Satisfactory</th>
<th>(2) Unsatisfactory</th>
<th>(1) Highly Unsatisfactory</th>
<th>Total</th>
<th>Valid Cases</th>
<th>Mean Score</th>
<th>Total Cases</th>
<th>Satisfactory or Better</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDC</td>
<td>15.2%</td>
<td>51.5%</td>
<td>27.3%</td>
<td>6.1%</td>
<td>100.0%</td>
<td>33</td>
<td>2.76</td>
<td>34.00</td>
<td>67%</td>
</tr>
<tr>
<td>MIC</td>
<td>19.2%</td>
<td>69.2%</td>
<td>3.8%</td>
<td>7.7%</td>
<td>100.0%</td>
<td>26</td>
<td>3.00</td>
<td>26.00</td>
<td>88%</td>
</tr>
</tbody>
</table>

3.1 Programs effectively address gender equality

<table>
<thead>
<tr>
<th>Country Type</th>
<th>(4) Highly Satisfactory</th>
<th>(3) Satisfactory</th>
<th>(2) Unsatisfactory</th>
<th>(1) Highly Unsatisfactory</th>
<th>Total</th>
<th>Valid Cases</th>
<th>Mean Score</th>
<th>Total Cases</th>
<th>Satisfactory or Better</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDC</td>
<td>8.0%</td>
<td>48.0%</td>
<td>32.0%</td>
<td>12.0%</td>
<td>100.0%</td>
<td>25</td>
<td>2.52</td>
<td>34</td>
<td>56%</td>
</tr>
<tr>
<td>MIC</td>
<td>6.7%</td>
<td>33.3%</td>
<td>53.3%</td>
<td>6.7%</td>
<td>100.0%</td>
<td>15</td>
<td>2.40</td>
<td>26</td>
<td>40%</td>
</tr>
</tbody>
</table>

4.2 Institutional and community capacity

<table>
<thead>
<tr>
<th>Country Type</th>
<th>(4) Highly Satisfactory</th>
<th>(3) Satisfactory</th>
<th>(2) Unsatisfactory</th>
<th>(1) Highly Unsatisfactory</th>
<th>Total</th>
<th>Valid Cases</th>
<th>Mean Score</th>
<th>Total Cases</th>
<th>Satisfactory or Better</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDC</td>
<td>9.1%</td>
<td>39.4%</td>
<td>33.3%</td>
<td>18.2%</td>
<td>100.0%</td>
<td>33</td>
<td>2.28</td>
<td>34</td>
<td>48%</td>
</tr>
<tr>
<td>MIC</td>
<td>16.0%</td>
<td>40.0%</td>
<td>40.0%</td>
<td>4.0%</td>
<td>100.0%</td>
<td>25</td>
<td>2.68</td>
<td>26</td>
<td>56%</td>
</tr>
</tbody>
</table>
### 5.1 Programs are cost/resource efficient

<table>
<thead>
<tr>
<th>Country Type</th>
<th>(4) Highly Satisfactory</th>
<th>(3) Satisfactory</th>
<th>(2) Unsatisfactory</th>
<th>(1) Highly Unsatisfactory</th>
<th>Total</th>
<th>Valid Cases</th>
<th>Mean Score</th>
<th>Total Cases</th>
<th>Satisfactory or Better</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDC</td>
<td>12.0%</td>
<td>56.0%</td>
<td>32.0%</td>
<td>.0%</td>
<td>100.0%</td>
<td>25</td>
<td>2.80</td>
<td>34</td>
<td>68%</td>
</tr>
<tr>
<td>MIC</td>
<td>4.5%</td>
<td>54.5%</td>
<td>40.9%</td>
<td>0%</td>
<td>100.0%</td>
<td>22</td>
<td>2.64</td>
<td>26</td>
<td>59%</td>
</tr>
</tbody>
</table>

### 5.2 Programs implemented/objectives achieved on time

<table>
<thead>
<tr>
<th>Country Type</th>
<th>(4) Highly Satisfactory</th>
<th>(3) Satisfactory</th>
<th>(2) Unsatisfactory</th>
<th>(1) Highly Unsatisfactory</th>
<th>Total</th>
<th>Valid Cases</th>
<th>Mean Score</th>
<th>Total Cases</th>
<th>Satisfactory or Better</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDC</td>
<td>10.0%</td>
<td>50.0%</td>
<td>20.0%</td>
<td>20.0%</td>
<td>100.0%</td>
<td>20</td>
<td>2.50</td>
<td>34</td>
<td>60%</td>
</tr>
<tr>
<td>MIC</td>
<td>0.0%</td>
<td>41.2%</td>
<td>41.2%</td>
<td>17.6%</td>
<td>100.0%</td>
<td>17</td>
<td>2.24</td>
<td>26</td>
<td>41%</td>
</tr>
</tbody>
</table>

### 6.4 Use of evaluation and monitoring to improve effectiveness

<table>
<thead>
<tr>
<th>Country Type</th>
<th>(4) Highly Satisfactory</th>
<th>(3) Satisfactory</th>
<th>(2) Unsatisfactory</th>
<th>(1) Highly Unsatisfactory</th>
<th>Total</th>
<th>Valid Cases</th>
<th>Mean Score</th>
<th>Total Cases</th>
<th>Satisfactory or Better</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDC</td>
<td>32.4%</td>
<td>17.6%</td>
<td>8.8%</td>
<td>41.2%</td>
<td>100.0%</td>
<td>34</td>
<td>2.41</td>
<td>34</td>
<td>50%</td>
</tr>
<tr>
<td>MIC</td>
<td>38.5%</td>
<td>26.9%</td>
<td>0.0%</td>
<td>34.6%</td>
<td>100.0%</td>
<td>26</td>
<td>2.69</td>
<td>26</td>
<td>65%</td>
</tr>
</tbody>
</table>

### 2. Comparison by Focus Area and Humanitarian Action

#### Criteria 1.1 Suited to the needs of target group members

<table>
<thead>
<tr>
<th>Focus Area</th>
<th>Highly Satisfactory</th>
<th>Unsatisfactory</th>
<th>Satisfactory</th>
<th>Highly Satisfactory</th>
<th>Not Addressed</th>
<th>Total</th>
<th>Addressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Young Child Survival and Development</td>
<td>1</td>
<td>0</td>
<td>11</td>
<td>8</td>
<td>1</td>
<td>21</td>
<td>20</td>
</tr>
<tr>
<td>2. Basic Education and Gender Equality</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>3. HIV/AIDS and Children</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4. Child Protection from Violence, Exploitation and Abuse</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>5. Policy Advocacy and Partnership</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>6. Humanitarian</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>7. Country Program</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>3</td>
<td>33</td>
<td>22</td>
<td>2</td>
<td>62</td>
<td>60</td>
</tr>
<tr>
<td>% of All Valid Cases Excluding Not Addressed</td>
<td>3%</td>
<td>5%</td>
<td>55%</td>
<td>37%</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Criteria 1.2 Programs aligned with national development goals

<table>
<thead>
<tr>
<th></th>
<th>Highly Unsatisfactory</th>
<th>Unsatisfactory</th>
<th>Satisfactory</th>
<th>Highly Satisfactory</th>
<th>Not Addressed</th>
<th>Total</th>
<th>Addressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Young Child Survival and Development</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>8</td>
<td>0</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>2. Basic Education and Gender Equality</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>4</td>
<td>0</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>3. HIV/AIDS and Children</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4. Child Protection from Violence, Exploitation and Abuse</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>5. Policy Advocacy and Partnership</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>6. Humanitarian</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>7. Country Program</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>0</strong></td>
<td><strong>3</strong></td>
<td><strong>31</strong></td>
<td><strong>27</strong></td>
<td><strong>1</strong></td>
<td><strong>62</strong></td>
<td><strong>61</strong></td>
</tr>
<tr>
<td>% of All Valid Cases Excluding Not Addressed</td>
<td>0%</td>
<td>5%</td>
<td>51%</td>
<td>44%</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Criteria 1.3 Effective partnerships

<table>
<thead>
<tr>
<th></th>
<th>Highly Unsatisfactory</th>
<th>Unsatisfactory</th>
<th>Satisfactory</th>
<th>Highly Satisfactory</th>
<th>Not Addressed</th>
<th>Total</th>
<th>Addressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Young Child Survival and Development</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>8</td>
<td>3</td>
<td>21</td>
<td>18</td>
</tr>
<tr>
<td>2. Basic Education and Gender Equality</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>3. HIV/AIDS and Children</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4. Child Protection from Violence, Exploitation and Abuse</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>5. Policy Advocacy and Partnership</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>6. Humanitarian</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>3</td>
<td>0</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>7. Country Program</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>3</strong></td>
<td><strong>4</strong></td>
<td><strong>32</strong></td>
<td><strong>17</strong></td>
<td><strong>6</strong></td>
<td><strong>62</strong></td>
<td><strong>56</strong></td>
</tr>
<tr>
<td>% of All Valid Cases Excluding Not Addressed</td>
<td>5%</td>
<td>7%</td>
<td>57%</td>
<td>30%</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*OECD DAC Network on Development Evaluation*
### Criteria 3.1 Effectively addresses gender equality

<table>
<thead>
<tr>
<th></th>
<th>Highly Unsatisfactory</th>
<th>Unsatisfactory</th>
<th>Satisfactory</th>
<th>Highly Satisfactory</th>
<th>Not Addressed</th>
<th>Total</th>
<th>Addressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Young Child Survival and Development</td>
<td>0</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>8</td>
<td>21</td>
<td>13</td>
</tr>
<tr>
<td>2. Basic Education and Gender Equality</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>3. HIV/AIDS and Children</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4. Child Protection from Violence, Exploitation and Abuse</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>5. Policy Advocacy and Partnership</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>6. Humanitarian</td>
<td>4</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>7. Country Program</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>6</strong></td>
<td><strong>16</strong></td>
<td><strong>17</strong></td>
<td><strong>3</strong></td>
<td><strong>20</strong></td>
<td><strong>62</strong></td>
<td><strong>42</strong></td>
</tr>
<tr>
<td>% of Total of All Valid Cases Excluding Not Addressed</td>
<td>14%</td>
<td>38%</td>
<td>40%</td>
<td>7%</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Criteria 6.1 Systems for evaluation effective

<table>
<thead>
<tr>
<th></th>
<th>Highly Unsatisfactory</th>
<th>Unsatisfactory</th>
<th>Satisfactory</th>
<th>Highly Satisfactory</th>
<th>Not Addressed</th>
<th>Total</th>
<th>Addressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Young Child Survival and Development</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>21</td>
<td>15</td>
</tr>
<tr>
<td>2. Basic Education and Gender Equality</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>3. HIV/AIDS and Children</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4. Child Protection from Violence, Exploitation and Abuse</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>5. Policy Advocacy and Partnership</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>6. Humanitarian</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td>7. Country Program</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>2</strong></td>
<td><strong>13</strong></td>
<td><strong>19</strong></td>
<td><strong>8</strong></td>
<td><strong>20</strong></td>
<td><strong>62</strong></td>
<td><strong>42</strong></td>
</tr>
<tr>
<td>% of All Valid Cases Excluding Not Addressed</td>
<td>5%</td>
<td>31%</td>
<td>45%</td>
<td>19%</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Criteria 6.2 Systems for monitoring and reporting results effective

<table>
<thead>
<tr>
<th>Category</th>
<th>Highly Unsatisfactory</th>
<th>Unsatisfactory</th>
<th>Satisfactory</th>
<th>Highly Satisfactory</th>
<th>Not Addressed</th>
<th>Total</th>
<th>Addressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Young Child Survival and Development</td>
<td>3</td>
<td>4</td>
<td>8</td>
<td>5</td>
<td>1</td>
<td>21</td>
<td>20</td>
</tr>
<tr>
<td>2. Basic Education and Gender Equality</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>3. HIV/AIDS and Children</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4. Child Protection from Violence, Exploitation and Abuse</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>5. Policy Advocacy and Partnership</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>6. Humanitarian</td>
<td>0</td>
<td>8</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>7. Country Program</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>7</td>
<td>29</td>
<td>16</td>
<td>7</td>
<td>3</td>
<td>62</td>
<td>59</td>
</tr>
<tr>
<td>% of All Valid Cases Excluding Not Addressed</td>
<td>12%</td>
<td>49%</td>
<td>27%</td>
<td>12%</td>
<td></td>
<td>100%</td>
<td></td>
</tr>
</tbody>
</table>
# Annex 7: UNICEF Staff Interviewed

<table>
<thead>
<tr>
<th>UNICEF Organizational Unit</th>
<th>Persons Interviewed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation Office</td>
<td>Colin Kirk</td>
</tr>
<tr>
<td></td>
<td>Samuel Bickel</td>
</tr>
<tr>
<td></td>
<td>Robert McCouch</td>
</tr>
<tr>
<td></td>
<td>Marco Segone</td>
</tr>
<tr>
<td></td>
<td>Abigail Taylor</td>
</tr>
<tr>
<td></td>
<td>Ashley Wax</td>
</tr>
<tr>
<td>Division of Financial and Administrative Management</td>
<td>Clair Jones</td>
</tr>
<tr>
<td>Division of Governance UN and Multilateral Affairs</td>
<td>Jean Dupraz</td>
</tr>
<tr>
<td></td>
<td>Bjorn Gillsater</td>
</tr>
<tr>
<td>Division of Policy and Strategy</td>
<td>Lakshmi Narasimhan Balaji</td>
</tr>
<tr>
<td></td>
<td>Etona Ekole</td>
</tr>
<tr>
<td></td>
<td>Robert Jenkins</td>
</tr>
<tr>
<td>Office of Emergency Programs</td>
<td>Genevieve Boutin</td>
</tr>
<tr>
<td></td>
<td>Akhil Iyer</td>
</tr>
<tr>
<td>Programme Division</td>
<td>Anju Malhotra</td>
</tr>
<tr>
<td></td>
<td>Christian Salazar</td>
</tr>
<tr>
<td>Public Sector Alliances and Resource Mobilization</td>
<td>Pieter Bult</td>
</tr>
<tr>
<td></td>
<td>Fernando Gutierrez-Eddy</td>
</tr>
</tbody>
</table>