

**EVALUATION OFFICE**

October 2016

# United Nations Children's Fund (UNICEF): Review of Development Effectiveness, 2012-2015

unite for  
children

unicef 

**UNITED NATIONS CHILDREN'S FUND (UNICEF):  
REVIEW OF DEVELOPMENT EFFECTIVENESS, 2012-2015**

Final Report

© United Nations Children's Fund, New York, 2016

United Nations Children's Fund  
Three United Nations Plaza  
New York, New York 10017

October 2016

The purpose of publishing evaluation reports produced by the UNICEF Evaluation Office is to fulfil a corporate commitment to transparency through the publication of all completed evaluations. The reports are designed to stimulate a free exchange of ideas among those interested in the topic and to assure those supporting the work of UNICEF that it rigorously examines its strategies, results, and overall effectiveness.

The contents of the report do not necessarily reflect the policies or views of UNICEF.

The text has not been edited to official publication standards and UNICEF accepts no responsibility for error.

The designations in this publication do not imply an opinion on the legal status of any country or territory, or of its authorities, or the delimitation of frontiers.

The copyright for this report is held by the United Nations Children's Fund. Permission is required to reprint/reproduce/photocopy or in any other way to cite or quote from this report in written form. UNICEF has a formal permission policy that requires a written request to be submitted. For non-commercial uses, the permission will normally be granted free of charge. Please write to the Evaluation Office at the address below to initiate a permission request.

For further information, please contact:

Evaluation Office  
United Nations Children's Fund  
Three United Nations Plaza  
New York, New York 10017  
[evalhelp@unicef.org](mailto:evalhelp@unicef.org)

# TABLE OF CONTENTS

<b>ACRONYMS</b>	<b>i</b>
<b>PREFACE</b>	<b>ii</b>
<b>EXECUTIVE SUMMARY</b>	<b>iv</b>
<b>RÉSUMÉ ANALYTIQUE</b>	<b>xiii</b>
<b>RESUMEN EJECUTIVO</b>	<b>xxiii</b>
<b>CHAPTER 1: INTRODUCTION</b>	<b>1</b>
1.1 Background	2
1.2 Purpose	2
1.3 Structure of the Report	3
1.4 Strategic Direction and Operations at UNICEF	3
1.5 Evaluation at UNICEF	6
1.6 The Evaluation Function at UNICEF	6
1.7 Results Monitoring and Reporting	8
<b>CHAPTER 2: METHODOLOGY</b>	<b>11</b>
2.1 Evaluation Population and Sample	12
2.2 Criteria	13
2.3 Review Process and Data Analysis	14
2.4 Limitations	14
<b>CHAPTER 3: FINDINGS OF UNICEF’S DEVELOPMENT EFFECTIVENESS</b>	<b>16</b>
3.1 Relevance	19
3.2 Achievement of Objectives	23
3.3 Cross-Cutting Themes	28
3.4 Sustainability	36
3.5 Efficiency	40
<b>CHAPTER 4: EVALUATION AND MONITORING FOR IMPROVED EFFECTIVENESS</b>	<b>45</b>
4.1 Coverage	46
4.2 Findings: Use of Evaluation	47
4.3 Findings: Monitoring for Improved Effectiveness	49
<b>CHAPTER 5: INNOVATION AND HUMAN RIGHTS-BASED APPROACHES</b>	<b>53</b>
5.1 Observations on Innovation	54
5.2 Observations on Human Rights-Based Approaches	55

## TABLE OF CONTENTS (CONT'D)

<b>CHAPTER 6: CONCLUSIONS AND RECOMMENDATIONS</b>	<b>57</b>
6.1 Conclusions	58
6.2 Recommendations	61
<b>ANNEXES</b>	<b>64</b>
Annex A – Development Effectiveness Criteria	64
Annex B – Evaluation Sample	65
Annex C – Methodology	70
Annex D – Comparison of Population and Sample	75
Annex E – Evaluation Quality Review Grid and Results	77
Annex F – Global Evaluations and Corporate Documents Reviewed	80
Annex G – List of Stakeholders Interviewed	82
Annex H – Comparison of Results by Country Type	83
Annex I – Comparison of Results of UNICEF DER 2016 and 2012	86
Annex J – Analysis Guide	89
Annex K – Terms of Reference	97
<b>FIGURES</b>	
Figure 1: Program Expenses by Output Area (USD 4,131 million) 2014	5
Figure 2: Program Expenses by Region (USD 4,131 million) 2014	5
Figure 3: Number of Evaluations Addressing Sub-Criteria for Relevance	19
Figure 4: Findings for Relevance	20
Figure 5: Number of Evaluations Addressing Sub-Criteria for Objectives Achievement	24
Figure 6: Findings for Objectives Achievement	25
Figure 7: Number of Evaluations Addressing Sub-Criteria for Cross-Cutting Themes	29
Figure 8: Findings on Effectiveness in Addressing Gender Equality	30
Figure 9: Number of Evaluations Addressing Sub-Criteria for Sustainability	36
Figure 10: Findings on Sustainability	37
Figure 11: Number of Evaluations Addressing Sub-Criteria for Efficiency	41
Figure 12: Findings on Efficiency	42
Figure 13: Number of Evaluations Addressing Sub-criteria for Evaluation and Monitoring	46
Figure 14: Findings for Using Evaluation and Monitoring	47

# ACRONYMS

<b>CEE/CIS</b>	Central and Eastern Europe and Commonwealth of Independent States	<b>ITN</b>	Insecticide Treated Bed Net
<b>CFS</b>	Child Friendly Schools	<b>LAC</b>	Latin America and the Caribbean
<b>CO</b>	Country Office	<b>MENA</b>	Middle-East Asia and North Africa
<b>CSG</b>	Child Support Grant	<b>MoRES</b>	Monitoring Results for Equity System
<b>CSN</b>	Country Strategy Note	<b>MTSP</b>	Medium-Term Strategic Plan 2006-2013
<b>DAC</b>	Development Assistance Committee	<b>NGO</b>	Non-Governmental Organization
<b>DER</b>	Development Effectiveness Review	<b>OECD</b>	The Organisation for Economic Cooperation and Development
<b>DRP</b>	Data, Research and Policy	<b>POC</b>	Point of Care
<b>EAPRO</b>	East Asia and the Pacific Regional Office	<b>QCPR</b>	Quadrennial Comprehensive Policy Review
<b>ECD</b>	Early Childhood Education	<b>RAM</b>	Results Assessment Module
<b>EMOPS</b>	Office of Emergency Programmes (UNICEF)	<b>RBM</b>	Results-based Management
<b>ESARO</b>	Eastern and Southern Africa Regional Office	<b>RO</b>	Regional Office
<b>EVALNET</b>	Network on Development Evaluation	<b>SDG</b>	Sustainable Development Goal
<b>FGM/C</b>	Female Genital Mutilation/Cutting	<b>UN</b>	The United Nations
<b>FRG</b>	Field Results Group	<b>UNDG</b>	United Nations Development Group
<b>GBV</b>	Gender Based Violence	<b>UNEG</b>	United Nations Evaluation Group
<b>GEROS</b>	Global Evaluation Report Oversight System	<b>UNICEF</b>	United Nations Children's Fund
<b>GGI</b>	Goss Gilroy Inc.	<b>USD</b>	United States Dollar
<b>HPM</b>	Humanitarian Performance Monitoring	<b>VISION</b>	Virtual Integrated System of Information
<b>IRI</b>	Interactive Radio Instruction	<b>WASH</b>	Water, Sanitation and Hygiene
		<b>WCARO</b>	Western and Central Africa Region

# PREFACE

This report presents a comprehensive synthesis of evaluation evidence responding, in essence, to two simple questions: how effective is UNICEF? And how do we know?

It is now more than a decade since a set of key principles of “aid effectiveness” were agreed and articulated in the Paris Declaration of 2005. However, the focus on the quality of aid and its impact was quickly overtaken by attention to “development effectiveness” and, in particular, a focus on the performance and contribution of multilateral organizations in the achievement of development results.

To address a perceived information gap around these issues, the evaluation network of the OECD-DAC developed an approach to assessing the development effectiveness of multilateral organizations using evidence from evaluations produced by the organizations themselves. Several “development effectiveness reviews” were prepared and published by the evaluation network, including a review of UNICEF’s development effectiveness based on evaluations completed in the period 2009-2011.

The findings of the UNICEF review undertaken in 2012 were broadly positive, although the review identified some significant shortcomings, notably with regard to issues of sustainability and attention to gender equality. Overall, despite certain limitations, the review was found to be a useful exercise and it fed into ongoing efforts to strengthen approaches to results-based management in the organization. A follow up review using the same methodology was included in the work plan of the Evaluation Office and duly commissioned in 2016. The present report sets out the results of this second development effectiveness review with the aim of providing UNICEF management, partners and stakeholders with a concise overview of UNICEF’s development effectiveness, based on evaluation evidence from the previous four years.

The conclusions of the review are again largely positive, despite the increasingly turbulent global development context. The report notes that UNICEF has given increased attention to equity and human rights-based programming, as well as greater commitment to innovation. However, certain areas are flagged for improvement. While the evaluations reviewed show increased attention to sustainability and gender equality, performance in these areas was still found to be mixed.

It is important to recognize the limitations of the approach taken by the review. While comprehensive and systematic, the methodology nevertheless depends on evaluation evidence about programmes which, in some cases, may have been designed a decade before. This is significant because UNICEF has continued to change and evolve while the organizational context has become ever more challenging. As acknowledged in the report, UNICEF has in recent years made significant efforts to strengthen “management for results” and to report

more systematically on its performance. It can be expected that the improving trend recorded in the present report will continue into the future. The recommendations presented in this report are intended to further support, strengthen and accelerate this progress.

The review was undertaken by a team fielded by the consulting company Goss Gilroy Inc. I would like to acknowledge the work of the team, composed of Ted Freeman (team leader) and Hubert Paulmer. I am also grateful to colleagues in UNICEF who shared their perspectives and experience through interviews and by reviewing drafts. Finally, I would like to thank colleagues in the Evaluation Office, Mathew Varghese, Senior Evaluation Specialist, served as the evaluation manager and was assisted by Laurence Reichel, Evaluation Specialist. Administrative support was provided by Geeta Dey, Celeste Lebowitz and Dalma Rivero.

The review asked how effectively UNICEF delivers against its development and humanitarian objectives, and how we know what has been achieved. I hope the present report serves to provide clear answers: not only a strong indication that UNICEF is indeed largely effective in the efforts it makes for children around the world, but also clear and credible evidence of this achievement, drawn from the body of evaluation work managed by the organization in recent years.

Colin Kirk  
Director  
Evaluation Office  
UNICEF

# EXECUTIVE SUMMARY

## Background

This report presents the results of the second review of the effectiveness of developmental and humanitarian programming supported by the United Nations Children’s Fund (UNICEF). The first was carried out in 2012 and 2013 and covered the results of UNICEF evaluations published from 2009 to 2011. The common approach and methodologies for both reviews (2012 and 2016) were developed under the guidance of the Development Assistance Committee (OECD DAC) Network on Development Evaluation. This review is based on the content of evaluation reports published by UNICEF between January 2012 and December 2015, supplemented by a review of corporate documents and consultations with UNICEF staff at headquarters in New York and in UNICEF regional offices (Annex F).

## Purpose

The purpose of this Development Effectiveness Review is to generate a body of credible information on a common set of criteria (as prescribed in the Guidance Document<sup>1</sup>) that provides an

insight into the development and humanitarian effectiveness of UNICEF. It involved conducting a systematic synthesis of information from a sample of available evaluations published from 2012 to 2015 as they relate to the criteria of development effectiveness. The review also is intended to inform decisions on issues such as adjustments to results-based management (RBM) systems and determining those effectiveness criteria which require adjustments in programming for more positive results.

## Approach and Methodology

The review was carried out from January to August 2016. It began with a preliminary review of UNICEF documents and the identification of the population of UNICEF evaluation reports found in UNICEF’s global data base and its Global Evaluation Oversight System. Interviews were also conducted with key personnel at UNICEF headquarters in New York City and with staff of each regional office (Annex F).

In consultation with evaluation staff, the review team identified a potential universe of 263 evaluation reports. From this universe, they drew a modified random sample of 73 evaluations which were then subjected to the team’s own quality review using an evaluation quality assessment template developed specifically for use in development effectiveness reviews (Annex D). Of these 73 reports, seven (9.6 percent) were eliminated because their quality scores were lower than the minimum required score (25 of 40 available points).

As a result, 66 evaluation reports were retained for systematic analysis, rating and coding of evaluation findings during the review. This compares to 62 evaluation reports subjected to the same process in 2012.

### Seven Outcome Areas

1. Health
2. HIV&AIDS
3. Water Sanitation and Hygiene (WASH)
4. Nutrition
5. Education
6. Child Protection
7. Social Inclusion

*UNICEF Strategic Plan 2014-2017*

<sup>1</sup> Assessing the Development Effectiveness of Multilateral Organizations: Guidance on the Methodological Approach (<http://www.oecd.org/dac/evaluation/dcdndep/50540172.pdf>)

Each evaluation report was reviewed to assess the findings it reported against six key development effectiveness criteria and 20 sub-criteria (Annex A). The findings for each sub-criterion were classified using a four-point scale that ranged from highly satisfactory on the positive side, to highly unsatisfactory on the negative. The review also identified the factors contributing to both positive and negative findings for each of the six criteria used to assess effectiveness as reported by the evaluations. The detailed results of this exercise in meta-synthesis of UNICEF evaluation report findings are presented in Chapters Three and Four.

### Assessment Criteria

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

### Coverage of the Assessment Criteria in Reviewed Evaluations

The review established ranges for assessing how well the 20 sub-criteria were covered in the 66 evaluations in the sample, based on the number of evaluations which adequately addressed each sub-criterion. Coverage could be rated strong, moderate or weak. The coverage achieved by the evaluations reviewed for the 2016 review was notably stronger than for the review in 2012.

### Coverage Ratings

**Strong** = covered in 46 or more evaluation reports

**Moderate** = covered in 31-45 evaluations

**Weak** = covered in less than 31 evaluations

### Limitations

The primary limitations of the review are related to sampling and to the retrospective nature of both evaluation and meta-synthesis as a methodology.

*Sampling Bias:* The evaluation sample, while it began with a stratified random sample, was adjusted to purposely improve representation of evaluations in the health sector and evaluations from countries in sub-Saharan Africa. This resulted in a purposive sample that was intended to be illustrative of the results of UNICEF evaluations published from 2012 to 2015. It is not a statistically valid random sample of UNICEF evaluations which can be used to calculate confidence intervals.

*Limits to Stratification:* The evaluation sample has been balanced to be largely reflective of the distribution of evaluation reports produced by UNICEF in relation to geographic area, year of publication, and outcome area of the strategic plan. However, the resulting strata of the sample are too small to allow for meaningful reporting of results in each sub-component (the different outcome areas for example).

*Assessing Effectiveness in Humanitarian Programming:* These limits to stratification of the sample have particular importance when considering the humanitarian dimension of UNICEF's work. Because evaluation reports with a direct focus on humanitarian programming accounted for only 2.7 percent of all UNICEF published evaluations during the review period, a very small number could be included in the review if the sample was to reflect the overall population. While directly

humanitarian focused evaluations were somewhat oversampled (they made up 6.1 percent of the sample), they still were too limited in number to allow any generalization of results to UNICEF's overall experience in humanitarian programming. It is important to note that the evaluation universe, and the sample, both include evaluations classified under other output areas, for example water and sanitation or health programming, which take place in a humanitarian or emergency contexts and so have potential implications for humanitarian programming. However, the actual review process of evaluations included in the sample did not provide enough evidence of findings on humanitarian operations to support conclusions on humanitarian programming.

*The Retrospective Nature of Meta-Synthesis:* Evaluation reports are, by definition, retrospective in that they provide ex-ante assessments of the effectiveness of programs. Many of the programs evaluated from 2012 to 2015 originated some years before they were evaluated. This means that the review must take care not to hold older programs accountable for new strategies or to generalize historical results to current practice.

While these limitations are real, none strongly undermine the utility of the results of the development effectiveness review process. By taking care in selecting the evaluation sample and in classifying and analyzing evaluation results, the review team has been able to replicate the 2012 process and provide a meaningful overview of the existing evaluative evidence of UNICEF's organizational effectiveness (and how it has changed in the intervening period).

### Recent Initiatives by UNICEF

The conclusions and recommendations presented here should not be read as indicating that UNICEF was not aware of or attempting to address the findings reported in evaluations published from 2012 to 2015. Recent and ongoing initiatives led by the Field Results Group and

aimed at strengthening programming (including Results-Based Management) reported to the review team include:

- Ongoing UNICEF RBM global training efforts (2015-2017) which have reached 1,500 staff in 2016 and a planned reach of 4,000 by the end of 2017. These efforts make use of a global cadre of trainers and a suite of face-to-face and e-learning modules.
- Program Strategy Notes – introduced in November 2015 and included in all new Country Programs. Staff report over 85 of the Strategy Notes have been developed and each includes a theory of change, a results framework and alignment of results and strategies with expected funding. A review of the quality of the Strategy Notes prepared in 2016 is currently underway.
- The implementation of the Country Office score card with a direct link from country office performance as measured by the score card to the personal performance assessment of the UNICEF Representative.

However, the terms of reference for the review and the methodology used did not include or permit any effort to assess the effectiveness of these initiatives in addressing the issues reported here.

### Conclusions

The detailed findings of the Development Effectiveness Review of UNICEF, 2016 are presented in detail in Chapters Three and Four of the report. This section presents the conclusions of the review.

### Relevance

1. As was the case in 2012, the 2016 Development Effectiveness Review **demonstrates that UNICEF programming remains highly relevant**. In particular, evaluations carried out from 2012 to 2015 reported that UNICEF supported **projects and programs**

**align well with national development priorities and programs.** While, evaluations report positive findings for relevance, they also identify **potential improvements** in the relevance of some programs through **engagement with key stakeholders during program development and through better links from needs assessment to targeted assistance.**

### Objectives Achievement

2. **UNICEF supported programs evaluated from 2012 to 2015 have been successful in attaining their stated objectives.** The evaluations reviewed reported a strongly positive and moderately improved level of objectives achievement by UNICEF supported programs when compared to results of the 2012 review. In particular, **many evaluation reports noted UNICEF has been successful in its upstream work in support of policy and strategy development,** and in supporting effective training and capacity development programs. These are both **areas of improvement since 2012:** when the previous Development Effectiveness Review identified weak capacity development work as a negative factor which sometimes limited objectives achievement. On the other hand, the reviewed evaluations also noted that some programs were not successful in reaching their stated objectives because of **weak program design, especially due to the absence of well-designed and clear theories of change.**

### Promoting Gender Equality

3. UNICEF's published **evaluations were much more likely to address issues of gender equality** in the period under review (2012-2015) than they did from 2009 to 2011, as reflected in the 2012 review. The higher level of attention paid to gender equality in evaluations published from 2012 to 2015 suggests that the **evaluation function at UNICEF has responded positively to the Strategic**

**Priority Action Plan** on gender equality (2010 to 2013), as well as the results of the 2012 development effectiveness review.

4. While coverage in evaluation reports published by UNICEF has improved since the 2012 review, **the actual findings on how effectively programs address gender equality presented in those reports continue to be very disappointing.** Almost exactly the same proportion of evaluations reported positive findings on gender equality for the 2016 (48 percent) and the 2012 review (47 percent). Now, as then, **less than half the UNICEF evaluation reports which address gender equality as cross-cutting theme, report positive findings.** This apparent lack of improvement comes despite the four-year Strategic Priority Action Plan (SPAP) from 2010 to 2013.
5. When UNICEF supported programs fail to adequately address gender equality as a cross-cutting theme, by far the most common reason given in evaluation reports is the **absence of a gender perspective and/or the non-application of gender analysis during program design and delivery.** This was also the most common factor impeding effectiveness in addressing gender equality as cited in the 2012 review. **These conclusions raise important questions about the appropriate balance between mainstreaming gender equality and the need for specific, targeted interventions aimed at addressing gender in UNICEF supported programs.**

### UNICEF's Focus on Equity

6. UNICEF evaluations published from 2012 to 2015 provide a strong indication that the **evaluation function has responded appropriately to the re-focus on equity** which began in 2010. A significant majority of the evaluations reviewed (54 of 66) directly addressed the issue of whether or not UNICEF programming effectively promotes improved equity for children and adolescents.

7. Results for this new sub-criterion on **effectiveness in improving equity were largely positive**. UNICEF supported programs were successful in addressing equity by targeting a diverse set of vulnerable or disadvantaged groups, including: poorest households by socio-economic strata; most distant, hard-to-reach rural districts; disadvantaged ethnic or tribal groups; migrant workers; children employed in urban settings and adolescent girls facing specific barriers to participation in school. It also seems clear, however, that effectively addressing equity requires UNICEF and its partners to clearly identify the disadvantaged as a target group, to map their location and to identify the barriers which impede their access to services or assets so that these can be addressed.
8. There remains considerable scope for improving the effectiveness of a significant portion of UNICEF supported programs in efforts to address inequality. For the most part this improvement would require UNICEF supported programs to **develop an explicit strategy to focus on benefiting marginalized individuals and communities; a strategy which is accompanied by criteria for participation and service access which provide a clear advantage to the marginalized**. This strategy to address marginalized groups would be needed for upstream policy work as well as for programs directed to beneficiaries.

### Sustainability

9. As with gender equality, **the evaluation function at UNICEF has improved the level of attention paid to issues of sustainability during the 2012 to 2015 period** when compared to evaluations carried out from 2009 to 2011.
10. Despite this improved level of coverage, evaluations reviewed in 2016 continue to report **challenging findings regarding sustainability**. There has been some improvement in findings relating to strengthened community and institutional capacity as well as

strengthening the policy environment since 2012. However, the basic question of whether or not the benefits of UNICEF programming will be sustained after program completion continues to generate negative findings in many evaluation reports (74 percent). In order for UNICEF supported projects and programs to improve the sustainability of results, it seems clear they will need to go beyond the noted improvements in capacity development and the enabling environment. **The evaluation reports point directly to the necessity for projects and programs to effectively address the need to plan for and ensure financial and program sustainability.**

### Efficiency

11. The majority of UNICEF programs evaluated since 2012 have been found to be **efficient in controlling costs and using resources**. They have also been successful **in implementing programs and achieving objectives on time**. However, as with the 2012 review, the evaluations point to the need for continuous attention to the cost and resource efficiency of UNICEF supported programs. They also highlight the need to strengthen systems and processes for monitoring program costs and results in order to manage for efficiency.

### Evaluation for Improved Effectiveness

12. The review noted continuing **improvements in UNICEF's use of evaluation as a function which can contribute to increasing the effectiveness of supported programs**, including continued improvements in the quality of evaluation reports and their utility as reflected in management responses. However, there remains a **need to ensure wider coverage and a systematic approach to setting priorities for program and project to be evaluated so that the most relevant programs are the subject of regular, high quality, evaluations**.

## Monitoring and Results-Based Management for Improved Effectiveness

13. While UNICEF has made **substantive efforts to improve the strength of monitoring systems, and the use of results-based management (RBM) methods** since 2012, this has **not resulted in recognizable improvements** in the strength of monitoring systems from the perspective of evaluations carried out from 2012 to 2015. This may reflect the relatively recent (2014) establishment of the Field Results Group, but it also indicates that effectively strengthening monitoring and RBM systems remains a serious challenge for UNICEF.
14. Strengthening results monitoring, reporting and management at UNICEF requires the agency to engage in a **range of measures including; capacity development** for UNICEF staff and, especially, **implementing partners**; improving the clarity of **theories of change** which underpin program designs; **establishing meaningful program targets** which are **monitored using systems implemented as early as possible in the program cycle**; and, provision of **technical assistance** to sustain data collection and results tracking by implementing partners. The mandate of the Field Results Group responds directly to many of these challenges.

## Innovation

15. Evaluations reviewed indicate that UNICEF has made **progress in fostering a commitment to innovation**, including a willingness to experiment with different models of program support and service delivery and to use emerging technologies in innovative ways. However, there are also indications that the **emphasis on innovation can be strengthened by developing a clear definition of innovative programming and focusing greater attention to programming innovations during the design of programs and services**. There is also scope for greater emphasis on documentation and taking proven innovations to scale.

## Human Rights-Based Approaches to Programming

16. UNICEF evaluations published from 2012 to 2015 provide evidence of **an increased interest in human rights-based approaches to programming**. While they indicate that many programs have made a serious attempt at a human rights-based approach they also note that some show a more rhetorical than real commitment. In particular, the evaluation reports noted **the need to address both the capacities of duty bearers to provide services and the ability of rights-holders to effectively demand those services**.

In summary, the findings of evaluations published by UNICEF in the 2012 to 2015 period, do show improvements in overall development effectiveness and in addressing key issues through evaluation. At the same time, the reviewed evaluations point to areas requiring further strengthening for improved effectiveness including improvements in programming for gender, equity and for sustainability.

## Recommendations

The following recommendations are based on the findings and conclusions presented in the report. They indicate that UNICEF should:

1. **Strengthen the capacity of UNICEF and its partners in order to improve program designs based on a strong theory of change**. While UNICEF supported programs generally achieve their stated objectives, the review also identifies scope for improvement, especially in the development of meaningful theories of change which adequately specify the links from program activities and inputs to results at output and outcome levels. This should include:
  - a) Capacity building of UNICEF staff and the staff of partners at country office level in the development of effective and technically sound theories of change during program design – including identifying key causal assumptions which can

be tested during evaluations. This will require improved, concrete guidance on developing effective theories of change;

- b) Capacity building and ongoing technical support to UNICEF staff and partners in the identification of meaningful program targets and appropriate indicators for tracking target achievement over time; and,
  - c) Linking work on strengthening program design with efforts to strengthen results-based monitoring and management systems (see recommendation seven).
2. **More consistently apply gender analysis methods and incorporate a gender perspective into program design and during all programming stages.** The lack of improvement in the level of positive findings for gender equality when comparing results of the 2012 and 2016 reviews points to the need to:
- a) Ensure that program design processes and approval systems fully incorporate gender analysis and a gender perspective prior to UNICEF commitment;
  - b) Strengthen the capacity of UNICEF staff and partners to make use of gender analysis tools during program design;
  - c) Strengthen the collection and use of gender disaggregated data in UNICEF programmes;
  - d) Engage with communities on the situation of girls and women during program design and incorporate organizations representing the interests of girls and women into stakeholder consultations;
  - e) Balance the use of a gender mainstreaming approach with targeted interventions specifically aimed at addressing gender inequalities (including the inclusion of specific, clear gender relevant objectives during the programme design phase); and,

- f) Continuing to strengthen the level of attention paid to gender equality as a cross-cutting issue in UNICEF led systems and processes for monitoring evaluation, to build on the progress made between the 2012 and 2016 reviews.

These measures are consistent with the more targeted approach adopted by UNICEF in the 2014 Gender Action Plan, especially given its strong emphasis on the need for programmes to identify and seek to address bottlenecks and barriers.

3. **More clearly define and identify target groups of disadvantaged women and children in all UNICEF supported programs** and clarify the disproportional benefits the program will bring to the target group. The results of the review indicate there is an opportunity to build on the current level of positive results by improving further the focus on equity. This should include:

- a) More consistent efforts to identify who are the most disadvantaged groups of women and children in the programming area during program design, including better mapping of their location in the communities being served;
- b) Better identification of barriers to participation and access to services for the identified groups of disadvantaged women and children so that programs specifically address them;
- c) Targeted criteria for program participation and access to services for the identified disadvantaged groups; including criteria which privilege their access and specifically anticipate a disproportionate positive result; and,
- d) Strengthened monitoring of participation, access and benefits accruing to members of the identified disadvantaged groups of women and children. This may include ensuring that selected programmes

address the need to strengthen national systems for data management to better address equity.

4. **More consistently and effectively address program and financial sustainability in the design of country programs and strategies.**

While there has been a notable improvement in the institutional sustainability of UNICEF supported programming, there is considerable evidence that a significant challenge to the sustainability of UNICEF supported projects and programs is the financial sustainability of the initiatives being supported. UNICEF should make every effort during program and project design and implementation to:

- a) Develop and implement mechanisms for generating income for supported projects which can be sustained after UNICEF support is completed;
- b) Where these are not appropriate, advocate effectively for host national, provincial and local government to provide necessary, line item budgets; and,
- c) Advocate for and help secure alternative sources of funding from multilateral agencies, bilateral donors or the private sector.

5. **Continue, and even increase efforts to improve the efficiency of UNICEF supported programs.**

As more evaluations address issues of cost and efficiency in UNICEF supported programs, there is clearly scope to strengthen the adequacy and level of attention paid to issues of efficiency and cost effectiveness in UNICEF programming during program design. The most urgent requirement is to strengthen program designs and program monitoring systems so that they track program costs over time and establish a credible link from unit costs to results. Otherwise, program managers cannot be expected to manage for improved efficiency. Efforts to improve efficiency

could also benefit from the development of a common, core methodology for assessing program efficiency at UNICEF.

6. **Improve the selection of programs for evaluation at UNICEF in order to ensure higher levels of program coverage and introduce more systematic decision making on priorities.**

While the recently approved Evaluation Policy recommends that all programmes should be evaluated at least once every five years, there is still scope for better linking the decision on which programs and projects should be evaluated first (given limited resources) to the requirements of both the overall agency Strategic Plan and the need to improve program effectiveness. This is especially important at the level of Country Offices, where the rationale for selecting programs and projects for evaluation is not always clear. It may also result in greater attention being given to the links between development and humanitarian programming being supported by UNICEF. The mechanism of the Costed Evaluation Plan (CEP) may represent an important channel for strengthening the process for prioritizing evaluation decisions at country office level.

7. **Continue and intensify efforts to strengthen results monitoring, reporting and management systems.**

This highlights the essential nature of the mandate of the Field Results Group in support of more effective results monitoring and management. In particular, there is a demonstrated need to:

- a) Develop the capacity of, in particular, UNICEF's implementing partners and their staff for designing, establishing and operating results monitoring and management systems;
- b) Improve the basic quality of project and program designs, especially as they relate to developing clear and realistic program theories of change (see recommendation one);

- c) Establish meaningful program results targets to monitor program performance against targets using systems implemented as near to the beginning of the program cycle as possible; and,
- d) Provide stronger technical assistance to sustain data collection and results monitoring and reporting by implementing partners.

8. **Develop and promote a more specific definition of innovation in UNICEF programming, one which encompasses programmatic as well as technology-based innovation.** Reviewed evaluations indicate that UNICEF supported programs frequently face opportunities to promote innovative methods and approaches for strengthening program design, implementation, technical support, service delivery mechanisms, and ongoing supervision and follow up. In many cases these may be as important or even more important than promoting the use of advanced technologies. There are also opportunities to strengthen the documentation of support to innovation so that national decision makers have better information when deciding whether or not to fund UNICEF supported innovations at national scale. Improving UNICEF's modelling and evidence gathering and reporting on innovative practices could also strengthen sustainability.

Taken together, these measures should help to ensure continued improvement in the findings and conclusions reported by UNICEF's evaluation function during the life of current and succeeding agency Strategic Plans.

# RÉSUMÉ ANALYTIQUE

## Contexte

Le présent rapport expose les résultats du deuxième examen de l'efficacité des programmes de développement et des programmes humanitaires soutenus par le Fonds des Nations Unies pour l'enfance (UNICEF). Le premier a été effectué en 2012 et 2013 et couvrait les résultats des évaluations de l'UNICEF publiées de 2009 à 2011. La démarche et les méthodes communes utilisées pour ces deux analyses (2012 et 2016) ont été mises au point avec les conseils du réseau sur l'évaluation en matière de développement du Comité d'aide au développement (CAD) de l'OCDE. Le présent examen s'appuie sur le contenu des rapports d'évaluation publiés par l'UNICEF entre janvier 2012 et décembre 2015, complété par un examen de documents institutionnels internes et des consultations avec le personnel du siège de l'UNICEF à New York et des bureaux régionaux de l'UNICEF (annexe F).

## Objectif

L'objectif du présent examen sur l'efficacité dans le domaine du développement est d'obtenir une masse d'informations crédibles sur un ensemble de critères commun (comme prescrit

par le Document d'orientation<sup>†</sup>) qui fournissent un aperçu sur l'efficacité des programmes de l'UNICEF dans les domaines du développement et de l'action humanitaire. Ce travail a consisté à réaliser une synthèse systématique d'un échantillon de rapports d'évaluations publiées de 2012 à 2015 qui sont en rapport avec les critères d'efficacité dans le domaine du développement. Cet examen est aussi destiné à informer les prises de décisions sur les ajustements à apporter au système de gestion axée sur les résultats, et à déterminer les modifications nécessaires sur les critères d'efficacité de la programmation afin d'obtenir des résultats plus positifs.

## Approche et méthodologie

Le présent examen a été conduit entre janvier et août 2016. Il a débuté par une analyse préliminaire de documents de l'UNICEF et par l'identification de l'ensemble des rapports d'évaluation de l'UNICEF qui se trouvent dans la base de données globale de l'organisation et dans son système globale de supervision des évaluations. Des entretiens ont également été organisés avec des membres clés du personnel du siège de l'UNICEF à New York et avec le personnel de chacun des bureaux régionaux (annexe F).

En consultation avec le personnel chargé des évaluations, l'équipe chargée de l'examen a identifié un ensemble potentiel de 263 rapports d'évaluation, et extrait de cette collection un échantillon aléatoire modifié de 73 rapports d'évaluations. Ceux-ci ont ensuite été soumis à l'analyse qualitative de l'équipe qui a utilisé une grille d'analyse mis au point spécifiquement pour les examens sur l'efficacité des actions dans le domaine du développement (annexe D) pour évaluer la qualité des rapports d'évaluations. Sept de ces 73 rapports (soit 9,6 %) ont été éliminés en raison d'une qualité inférieure au seuil minimum requis (25 points sur 40 points).

### Sept domaines de réalisations

1. Santé
2. VIH & SIDA
3. Eau, assainissement et hygiène (EAH)
4. Nutrition
5. Éducation
6. Protection de l'enfance
7. Inclusion sociale

*Plan stratégique 2014-2017 de l'UNICEF*

<sup>†</sup> Assessing the Development Effectiveness of Multilateral Organizations: Guidance on the Methodological Approach (<http://www.oecd.org/dac/evaluation/dcdndep/50540172.pdf>)

Par conséquent, 66 rapports d'évaluation ont été retenus pour faire partie de l'exercice. Par comparaison, 62 rapports d'évaluation avaient été soumis à la même procédure en 2012.

Chacun de ces rapports d'évaluation a été examiné afin de jauger les résultats qu'il présentait en fonction de six critères clés d'efficacité, ainsi que de 20 sous-critères (annexe A). Les constatations pour chaque sous-critère ont été classées sur une échelle de quatre notes allant de « très satisfaisant » pour l'aspect positif à « très insatisfaisant » pour l'aspect négatif. Cet examen a également identifié les facteurs contribuant à des constatations positives et négatives pour chacun des six critères utilisés pour mesurer l'efficacité selon les rapports d'évaluation. Les résultats détaillés de cet exercice de méta-synthèse des constatations des rapports d'évaluation de l'UNICEF sont présentés dans les chapitres 3 et 4.

### Critères d'évaluation

1. Pertinence de l'intervention ;
2. Degré de réalisation des objectifs humanitaires et de développement ;
3. Thèmes transversaux (promotion de l'égalité des sexes et de l'équité) ;
4. Durabilité des résultats et des bénéfices issues des interventions ;
5. Efficience ;
6. Utilisation des outils de suivi-évaluation pour améliorer l'efficacité.

### Couverture des critères d'évaluation dans les rapports examinés

L'examen réalisé a défini des fourchettes permettant de juger dans quelle mesure les 20 sous-critères étaient couverts dans les 66 évaluations qui constituent l'échantillon, en prenant pour base le nombre d'évaluations qui tenaient adéquatement compte de chacun de ces sous-critères. Cette couverture pouvait être jugée étendue, modérée ou faible. La

couverture réalisée par les rapports d'évaluation examinés pour l'examen 2016 était nettement meilleure que celle de l'examen de 2012.

### Échelle de couverture

**Étendue** = couverts dans 46 rapports d'évaluation ou plus

**Modérée** = couverts dans 31-45 évaluations

**Faible** = couverts dans moins de 31 évaluations

### Limitations

Les principales limitations du présent exercice se rapportent à l'échantillonnage et à la nature rétrospective de l'évaluation comme de la méta-synthèse en tant que méthodologie.

*Biais d'échantillonnage* : l'échantillon d'évaluation, constitué originellement d'un échantillon aléatoire stratifié, a été volontairement ajusté afin d'améliorer la représentation des évaluations concernant le secteur de la santé et de celles provenant des pays d'Afrique subsaharienne. Ceci produit un échantillon raisonné dans l'intention d'illustrer les résultats des évaluations de l'UNICEF publiées de 2012 à 2015. Il ne constitue pas un échantillon aléatoire des évaluations de l'UNICEF statistiquement valide pouvant être utilisé pour calculer des intervalles de confiance.

*Limites de la stratification* : l'échantillon d'évaluation a été équilibré afin de refléter globalement la distribution des rapports d'évaluation produits par l'UNICEF en termes géographiques, d'année de publication et de domaine de réalisation défini par le plan stratégique. La stratification de l'échantillon qui en résulte est cependant trop réduite pour obtenir des résultats significatifs pour chacune de ses sous-composantes (par exemple pour les différents domaines de réalisation).

*Évaluation de l'efficacité dans la programmation humanitaire* : Ces limites de la stratification de l'échantillon ont une importance particulière quand on envisage la dimension humanitaire

du travail de l'UNICEF. Étant donné que les rapports d'évaluation directement axés sur les programmes humanitaires ne représentaient que 2.7 % de toutes les évaluations publiées par l'UNICEF au cours de la période d'examen, seul un petit nombre d'entre elles ont pu être comprises dans l'échantillon pour que celui-ci reflète adéquatement l'ensemble de tous ces rapports. Bien que les évaluations directement axées sur l'humanitaire aient été quelque peu surreprésentées dans l'échantillon (elles constituent 6,1 % de celui-ci), leur nombre était encore trop limité pour permettre une généralisation des résultats concernant l'expérience globale de l'UNICEF dans le domaine des programmes humanitaires. Il est important de noter que l'ensemble des évaluations, aussi bien que l'échantillon, comprennent des évaluations classées sous les rubriques d'autres domaines de réalisation, par exemple, eau et assainissement, ou programmes de santé, qui se déroulent dans le contexte d'interventions humanitaires ou d'interventions d'urgence, ayant donc ainsi des implications potentielles pour les programmes humanitaires. Cependant le processus d'examen des évaluations incluses dans l'échantillon n'a pas fourni suffisamment de données factuelles sur les résultats des opérations humanitaires pour tirer des conclusions concernant les programmes humanitaires.

*Nature rétrospective de la méta-synthèse* : les rapports d'évaluation sont par définition rétrospectifs en ce qu'ils ne fournissent que des évaluations ex ante de l'efficacité des programmes. Un grand nombre des programmes évalués de 2012 à 2015 remontaient à plusieurs années avant ces évaluations. Ceci signifie que la présente analyse doit prendre soin d'éviter de rendre des programmes plus anciens redevables des nouvelles stratégies ou de généraliser des résultats historiques aux pratiques actuelles.

Bien que ces limitations soient bien réelles, aucune n'affaiblit fortement l'utilité des résultats du processus d'examen de l'efficacité dans le domaine du développement. En procédant avec soin à la sélection de l'échantillon d'évaluation ainsi qu'à la classification et à l'analyse

des résultats d'évaluation, l'équipe chargée de cet examen a été à même de reprendre les procédures de 2012 et de fournir une bonne vue d'ensemble des constatations tirées des évaluations existantes relatives à l'efficacité de l'UNICEF en tant qu'organisation (et de la manière dont celle-ci a changé au cours de la période intermédiaire).

### Les récentes initiatives de l'UNICEF

Les conclusions et les recommandations exposées dans le présent document ne doivent pas être interprétées comme si l'UNICEF n'eut pas intégré ou tenté de tenir compte des résultats des évaluations publiées de 2012 à 2015. Les initiatives récentes et en cours lancées sous la direction du Groupe des résultats de terrain (*Field Results Group*) et visant à renforcer la programmation (y compris la gestion axée sur les résultats) qui ont été rapportées à l'équipe de rédaction du présent examen comprennent entre autres :

- Les efforts entrepris actuellement pour la formation à la gestion axée sur les résultats au sein de l'UNICEF (2015-2017) ; en 2016, la formation a atteint 1500 membres du personnel et il est prévu d'en former 4000 d'ici la fin de 2017. Cette initiative a recours à des formateurs répartis à travers le monde et à un ensemble de modules de formation présentielle et de formation en ligne.
- Les notes de stratégie programmatique – introduites en novembre 2015 et incluses dans tous les nouveaux programmes de pays. Le personnel fait état de l'élaboration de plus de 85 notes de stratégie programmatique, chacune comportant une théorie du changement, un cadre de résultats et un alignement des résultats et des stratégies ainsi que les financements attendus. Un examen de la qualité des notes de stratégie rédigées en 2016 est actuellement en cours.
- La mise en œuvre du tableau de bord de bureau de pays, avec un lien direct entre les résultats obtenus par le bureau de pays

tels que mesurés par ce tableau de bord et l'évaluation personnelle du Représentant de l'UNICEF.

Cependant, les termes de référence du présent examen et la méthodologie utilisée ne comportait pas d'éléments permettant de juger de l'efficacité de ces initiatives par rapport aux questions mentionnées dans le présent rapport.

## Conclusions

Les résultats obtenus par l'Examen de l'efficacité de l'UNICEF dans le domaine du développement de 2016 sont présentés de manière détaillée dans les chapitres 3 et 4 du présent rapport. Les paragraphes ci-dessous présentent les conclusions de l'examen.

### Pertinence

1. Comme cela avait été le cas en 2012, l'Examen 2016 de l'efficacité de l'UNICEF dans le domaine du développement **démontre que les programmes de l'UNICEF conservent une forte pertinence**. En particulier, les évaluations menées de 2012 à 2015 constatent que **les programmes et les projets qui bénéficient du soutien de l'UNICEF sont bien alignés sur les priorités et les programmes de développement nationaux**. Tout en faisant état de résultats positifs en termes de pertinence, les évaluations identifient également **des améliorations potentielles** de la pertinence de certains programmes à travers **une collaboration avec les parties prenantes clés au cours de l'élaboration du programme et par une meilleure liaison entre l'identification des besoins et l'assistance ciblée**.

### Réalisation des objectifs

2. **Les programmes soutenus par l'UNICEF qui ont été évalués de 2012 à 2015 ont réussi à atteindre les objectifs fixés**. Les évaluations examinées font état d'un degré de réalisation des objectifs fixés fortement positif. Ce qui constitue une amélioration modérée par rapport à l'examen de 2012

pour ces programmes bénéficiant du soutien de l'UNICEF. En particulier, **de nombreux rapports d'évaluation ont noté que l'UNICEF a mené avec succès son travail visant à soutenir l'élaboration des politiques et des stratégies**, ainsi que ses efforts favorisant des programmes de développement des capacités et de formation. **Ces deux domaines présentent des améliorations** par rapport à l'examen précédent qui avait identifié un travail de développement des capacités insuffisant comme facteur négatif limitant parfois la réalisation des objectifs fixés. D'un autre côté, les évaluations examinées ont aussi noté que certains programmes n'avaient pas réussi à atteindre les objectifs définis en raison de la **faiblesse du travail de conception des programmes résultant plus spécialement de l'absence de théories du changement claires et bien conçues**.

### Promotion de l'égalité des sexes

3. **Les évaluations** publiées par l'UNICEF **avaient une plus forte tendance à traiter des questions de l'égalité des sexes** au cours de la période examinée (2012-2015) que pendant celle de 2009 à 2011, comme le reflétait l'examen de 2012. La plus grande attention portée à l'égalité des sexes dans les évaluations publiées de 2012 à 2015 suggère que **la fonction d'évaluation de l'UNICEF a réagi positivement par rapport au Plan d'action stratégique prioritaire de l'UNICEF en faveur de l'égalité des sexes** (2010 à 2013) ainsi qu'aux résultats de l'examen 2012 sur l'efficacité de l'UNICEF dans le domaine du développement.
4. Si la couverture dans les rapports d'évaluation publiés par l'UNICEF s'est améliorée depuis l'examen de 2012, **les constatations sur l'efficacité des programmes concernant l'égalité des sexes que présentent ces rapports sont toujours très décevantes**. Le pourcentage des évaluations rapportant des résultats positifs dans le domaine de l'égalité des sexes pour 2016 est presque

exactement le même (48 %) que celui de l'examen de 2012 (47 %). Aujourd'hui comme à cette époque, **moins de la moitié des rapports d'évaluation de l'UNICEF qui traitent de l'égalité des sexes à titre de thème transversal font état de résultats positifs**. Cette apparente absence d'amélioration se manifeste en dépit du plan d'action stratégique prioritaire de 2010 à 2013.

5. Quand les programmes soutenus par l'UNICEF ne réussissent pas à traiter de manière adéquate la question de l'égalité des sexes comme thème transversal, la raison la plus communément avancée dans les rapports d'évaluation est l'absence de perspective intégrant la dimension du genre et/ou la non-application d'une analyse fondée sur le genre dans la conception des programmes comme dans leur mise en œuvre. C'est aussi le facteur le plus communément cité dans l'examen de 2012 comme obstacle à l'efficacité dans le traitement de l'égalité des sexes. **Ces conclusions soulèvent d'importantes questions sur la manière d'obtenir un équilibre adéquat entre l'intégration d'une démarche soucieuse d'égalité entre les sexes et la nécessité d'interventions spécifiques et ciblées visant à traiter de la question du genre dans les programmes soutenus par l'UNICEF.**

#### Interventions de l'UNICEF axées sur l'équité

6. Les évaluations de l'UNICEF publiées de 2012 à 2015 indiquent clairement que **la fonction d'évaluation a réagi de manière appropriée à la réorientation sur la question de l'équité** amorcée en 2010. Une nette majorité des évaluations examinées (54 sur 66) traitaient directement de la question de savoir si les programmes de l'UNICEF faisaient ou non une promotion efficace de l'équité pour les enfants et les adolescents.
7. Les résultats pour ce nouveau sous-critère de mesure de **l'efficacité dans l'amélioration de l'équité ont été largement positifs**. Les programmes soutenus par l'UNICEF ont

traité avec succès la question de l'équité en ciblant un ensemble diversifié de groupes vulnérables ou défavorisés, notamment : les ménages les plus pauvres par catégorie socio-économique, les populations des districts ruraux les plus éloignés et les plus difficiles à atteindre, les groupes ethniques ou tribaux défavorisés, les travailleurs migrants, les enfants employés en milieu urbain et les adolescentes faisant face à des obstacles spécifiques à leur scolarisation. Il semble cependant également clair que **traiter efficacement l'équité exige que l'UNICEF et ses partenaires identifient clairement les populations défavorisées comme groupe cible, localiser les endroits où elles sont présentes et les obstacles qui entravent leur accès aux services ou aux biens afin d'intervenir sur ces facteurs.**

8. Il reste un travail d'une ampleur considérable à entreprendre pour améliorer l'efficacité d'une part importante des programmes soutenus par l'UNICEF dans le cadre des efforts pour éliminer les inégalités. Ces améliorations exigent entre autres pour les programmes soutenus par l'UNICEF de **définir une stratégie explicite visant avant tout à procurer des avantages aux individus et aux communautés marginalisés, une stratégie accompagnée de critères sur la participation et l'accès aux services qui apportent un clair avantage à ces individus et communautés marginalisés**. Cette stratégie visant les groupes marginalisés sera requise dans le travail à faire sur les politiques ainsi que sur les programmes s'adressant aux bénéficiaires.

#### Durabilité

9. Comme dans le cas de l'égalité des sexes, **la fonction d'évaluation de l'UNICEF a amélioré le degré d'attention portée aux questions de durabilité au cours de la période de 2012 à 2015** par comparaison avec les évaluations menées de 2009 à 2011.

10. En dépit de cette amélioration du degré de couverture, les évaluations examinées en 2016 continuent à faire état de **constatations problématiques en ce qui concerne la question de la durabilité**. Les constatations concernant le renforcement des capacités communautaires et institutionnelles se sont améliorées, il en est de même pour le renforcement de l'environnement des politiques depuis 2012. Cependant, la question fondamentale de savoir si les résultats obtenus par les programmes de l'UNICEF subsisteront après la fin de ceux-ci continue à susciter des constatations négatives dans de nombreux rapports d'évaluation (74 %). Pour que les projets et les programmes soutenus par l'UNICEF améliorent la durabilité de leurs résultats, il semble clair qu'ils devront aller au-delà des améliorations observées pour ce qui est du renforcement des capacités et de l'environnement favorable. **Les rapports d'évaluation soulignent directement la nécessité pour les projets et programmes d'anticiper efficacement la question de la durabilité des interventions dans la planification et le financement afin de garantir la durabilité des programmes et de leur financement.**

#### Efficiences

11. La majorité des programmes de l'UNICEF évalués depuis 2012 se sont révélés **efficaces du point de vue du contrôle des coûts et de l'usage des ressources**. Ils ont également été **mis en œuvre et atteint leurs objectifs dans les délais impartis**. Cependant, comme dans le cas de l'examen de 2012, ces évaluations soulignent le besoin de surveiller en permanence les coûts et l'efficacité de l'utilisation des ressources des programmes soutenus par l'UNICEF. Elles soulignent aussi le besoin de renforcer les dispositifs et les procédures de suivi des coûts et des résultats des programmes afin de gérer pour l'efficacité des programmes.
- L'évaluation au service d'une meilleure efficacité**
12. L'analyse a noté des **améliorations continues de l'usage que fait l'UNICEF de l'évaluation comme fonction pouvant contribuer à renforcer l'efficacité des programmes qu'il soutient**, y compris des améliorations de la qualité des rapports d'évaluation et de leur utilité telles que reflétées par les réponses de la direction. Il persiste cependant **un besoin d'assurer une couverture plus large et de définir de manière systématique les priorités des programmes et des projets à évaluer, de façon à ce que les programmes les plus pertinents soient l'objet d'évaluations régulières et de haute qualité.**
- Le suivi et la gestion axée sur les résultats au service d'une meilleure efficacité**
13. Bien que l'UNICEF ait fait depuis 2012 **des efforts substantiels pour renforcer les systèmes de suivi et l'utilisation des méthodes de gestion axée sur les résultats, ceci n'a pas produit de résultats identifiables** en ce qui concerne la robustesse de ces systèmes de suivi du point de vue des évaluations réalisées de 2012 à 2015. La raison pourrait en être la mise en place relativement récente (2014) du Groupe des résultats de terrain (*Field Results Group*), mais cela indique également qu'un renforcement effectif des dispositifs de suivi et de gestion axée sur les résultats continue à représenter un sérieux défi pour l'UNICEF.
14. Le renforcement du suivi des résultats, du rapportage et de la gestion à l'UNICEF exige que l'organisation prenne **une série de mesures incluant : le développement des capacités** pour le personnel de l'UNICEF, et en particulier pour **les partenaires de mise en œuvre**, l'amélioration de la clarté des **théories du changement** qui sous-tendent la conception des programmes, **la définition pour les programmes de cibles judicieuses dont le suivi se fait grâce à des dispositifs mis en place aussitôt que possible dans le cycle de programme**, et la fourniture d'une **assistance technique** pour soutenir la

collecte de données et le suivi des résultats par les partenaires de mise en œuvre. Le mandat du Groupe des résultats de terrain (*Field Results Group*) répond directement à un grand nombre de ces défis.

### L'innovation

15. Les évaluations examinées indiquent que l'UNICEF a fait **des progrès pour ce qui est de favoriser un engagement en faveur de l'innovation**, y compris en démontrant sa volonté d'expérimenter différents modèles de soutien aux programmes et à la prestation de services et d'utiliser les technologies émergentes de manière novatrice. Il existe cependant aussi des indications selon lesquelles **l'accent mis sur l'innovation peut être renforcé grâce à une définition claire de l'innovation dans les programmes, et en insistant davantage sur l'attention à porter à ce facteur au cours de la conception des programmes et des services**. Il reste du travail à faire pour mettre davantage l'accent sur la documentation et de mettre à grande échelle les innovations ayant fait leurs preuves.

### Programmation axée sur l'approche des droits de l'homme

16. Les évaluations de l'UNICEF publiées de 2012 à 2015 prouvent l'existence d'**un intérêt accru pour une programmation axée sur les droits de l'homme**. Tout en indiquant que de nombreux programmes ont sérieusement tenté de définir une démarche axée sur les droits de l'homme, elles notent également que certains manifestent un engagement plus rhétorique qu'un véritable engagement. Les rapports d'évaluation notent en particulier **le besoin de traiter aussi bien la question des capacités des débiteurs d'obligation à fournir des services que celle de la possibilité pour les détenteurs de droits d'exiger effectivement ces services**.

En résumé, les constatations des évaluations publiées par l'UNICEF dans la période 2012-2015 indiquent bien des améliorations de l'efficacité

globale des interventions de l'UNICEF dans le domaine du développement et dans le traitement de questions clés à travers l'évaluation. Les évaluations analysées signalent des domaines qui exigent un renforcement supplémentaire pour obtenir une amélioration de l'efficacité de l'action engagée, y compris par des améliorations dans la programmation en faveur de l'égalité des sexes, de l'équité et de la durabilité.

### Recommandations

Les recommandations suivantes sont fondées sur les constatations et les conclusions contenues dans le présent rapport. Elles indiquent que l'UNICEF devrait :

1. **Renforcer les capacités de l'UNICEF et de ses partenaires afin d'améliorer la conception des programmes fondés sur une solide théorie du changement**. Si les programmes soutenus par l'UNICEF atteignent généralement les objectifs fixés, le présent examen a également identifié des domaines à améliorer, spécialement en ce qui concerne la mise au point de théories du changement utiles qui spécifient adéquatement les liaisons entre les activités et les ressources, les résultats obtenus au niveau des produits et des réalisations. Il faudrait inclure :
  - a) Renforcer les capacités du personnel de l'UNICEF et du personnel des partenaires au niveau des bureaux de pays dans le but de mettre au point des théories du changement efficaces et techniquement valides au cours de la phase de conception des programmes – y compris par l'identification des hypothèses causales clés susceptibles d'être testées en cours d'évaluation. Ceci exigera un travail d'orientation amélioré et concret sur la mise au point de ces théories du changement ;
  - b) Renforcer les capacités et fournir un soutien technique continu du personnel de l'UNICEF et des partenaires dans l'identification de cibles programmatiques

judicieuses et d'indicateurs appropriés permettant de suivre la progression vers les cibles ;

- c) Lier le renforcement de conception des programmes à des efforts pour consolider les systèmes de suivi et de gestion axés sur les résultats (voir recommandation 7).

**2. Appliquer de manière plus cohérente des méthodes d'analyse basées sur l'égalité des sexes et intégrer une perspective basée sur l'égalité des sexes dans la conception des programmes et à toutes les étapes de programmation.** L'absence d'amélioration au niveau de résultats positifs dans le domaine sur l'égalité des sexes quand on compare les résultats des examens de 2012 et de 2016 indique qu'il est nécessaire de :

- a) S'assurer, avant tout engagement de l'UNICEF, que les processus de conception des programmes et les dispositifs d'approbation intègrent pleinement une analyse des questions basées sur l'égalité des sexes et une perspective qui en tient compte ;
- b) Renforcer la capacité du personnel de l'UNICEF et des partenaires à faire usage des outils d'analyse de l'égalité des sexes pendant la phase de conception des programmes ;
- c) Renforcer la collecte et l'utilisation de données désagrégées par sexe pour les programmes de l'UNICEF ;
- d) Engager le dialogue avec les communautés sur la situation des filles et des femmes pendant la phase de conception des programmes et intégrer dans les consultations auprès des parties prenantes des organisations représentant les intérêts des filles et des femmes ;
- e) Équilibrer l'intégration d'une démarche d'égalité entre les sexes avec des interventions ciblées visant à traiter les inégalités existantes (y compris en fixant des

objectifs spécifiques, clairs et pertinents à ce sujet pendant la phase de conception des programmes) ;

- f) Continuer à renforcer l'attention portée à l'égalité entre les sexes à titre de question transversale à intégrer dans les dispositifs et les processus de suivi de l'évaluation afin de faire avancer les progrès réalisés entre les examens de 2012 et de 2016.

Ces mesures sont en accord avec la démarche plus ciblée adoptée par l'UNICEF dans son Plan d'action pour l'égalité des sexes de 2014-2017, compte tenu en particulier de l'accent mis sur le besoin d'identifier et de chercher à surmonter les barrières et les goulots d'étranglement.

**3. Définir et identifier plus clairement les groupes cibles de femmes et d'enfants défavorisés dans tous les programmes soutenus par l'UNICEF** et clarifier les avantages disproportionnés qu'un programme apportera au groupe ciblé. Les résultats du présent examen indiquent qu'il existe une possibilité de consolider le niveau actuel des constatations positives en insistant encore plus sur l'accent placé sur l'équité. Ce qui devrait inclure :

- a) Des efforts plus cohérents pour identifier les groupes de femmes et d'enfants les plus désavantagés dans le domaine d'activité des programmes durant leur phase de conception, y compris par un meilleur recensement géographique de leur présence dans les communautés servies ;
- b) Une meilleure identification des obstacles à la participation et à l'accès aux services rencontrés par les groupes identifiés de femmes et d'enfants désavantagés, de manière à ce que les programmes répondent spécifiquement à leurs problèmes ;
- c) Des critères ciblés applicables à la participation et à l'accès aux services des groupes désavantagés qui ont été identifiés, y compris des critères qui

privilégient leur accès et prévoient spécifiquement des résultats disproportionnellement positifs ;

- d) Le renforcement du suivi de la participation, de l'accès et des avantages dont bénéficient les membres de ces groupes de femmes et d'enfants désavantagés qui ont été identifiés. Ceci peut impliquer d'assurer que des programmes sélectionnés répondent au besoin de renforcer les systèmes nationaux de gestion des données afin de mieux traiter la question de l'équité.
4. **Intégrer de manière plus cohérente et plus efficace la durabilité programmatique et financière dans la conception des programmes de pays et des stratégies.** Bien qu'il y ait eu une amélioration notable dans la durabilité institutionnelle des projets et des programmes soutenus par l'UNICEF, un nombre considérable de données factuelles indique que la durabilité financière des initiatives que l'organisation appuie constitue un problème important pour la durabilité de ces projets et de ces programmes. L'UNICEF devra faire tous les efforts possibles au cours de la conception et de la mise en œuvre de ces programmes et de ces projets pour :
- a) Élaborer et mettre en œuvre des mécanismes destinés à générer des revenus pour les projets soutenus afin qu'ils puissent être maintenus lorsque prend fin l'appui apporté par l'UNICEF ;
- b) Là où ceux-ci ne sont pas appropriés, faire un plaidoyer auprès des autorités nationales, provinciales et locales pour fournir un budget nécessaire, par objets de dépense ;
- c) Faire un plaidoyer auprès des agences multilatérales, des bailleurs de fonds bilatéraux ou du secteur privé pour obtenir de nouvelles sources de financement.
5. **Poursuivre, et même renforcer les efforts pour améliorer l'efficacité des programmes soutenus par l'UNICEF.** Dans la mesure où plus d'évaluations traitent des questions de coûts et d'efficacité dans les programmes soutenus par l'UNICEF, il existe un travail important à accomplir pour renforcer l'adéquation et le niveau d'attention portée aux questions d'efficacité et du rapport coûts-efficacité dans la conception des programmes de l'UNICEF. L'exigence la plus urgente est de renforcer la conception des programmes et leurs dispositifs de suivi pour leur permettre de suivre l'évolution des coûts de programmation et d'établir une liaison claire entre les coûts unitaires et les résultats, sinon il sera difficile à ce que les gestionnaires de programme assurent l'amélioration de cette efficacité. Les efforts d'amélioration de l'efficacité pourraient également bénéficier de la définition d'une méthodologie de base commune pour évaluer l'efficacité des programmes de l'UNICEF.
6. **Améliorer la sélection des programmes de l'UNICEF à évaluer afin d'assurer une meilleure couverture des programmes et d'introduire une prise de décision plus systématique concernant les priorités.** Bien que la politique d'évaluation récemment approuvée recommande que tous les programmes soient évalués au moins une fois tous les cinq ans, il y a besoin d'améliorer les liens entre la prise de décision sur quels programmes et projets doivent être évalués en premier (étant donné les ressources limitées) et les exigences du plan stratégique d'UNICEF ainsi que le besoin d'améliorer l'efficacité des programmes. Ceci est particulièrement important au niveau des bureaux de pays où la raison pour sélectionner des programmes et des projets à évaluer n'est pas toujours claire. Cela pourrait aussi conduire à accorder plus d'attention aux liens entre les programmes de développement et les programmes humanitaires soutenus par l'UNICEF. Le mécanisme de plan d'évaluation chiffré peut représenter

un important moyen de renforcer le processus pour déterminer les priorités relatives aux décisions d'évaluation prises au niveau des bureaux de pays.

**7. Poursuivre et intensifier les efforts pour renforcer le suivi des résultats, le rapportage et les systèmes de gestion.** Cette recommandation souligne la nature essentielle du mandat du Groupe des résultats de terrain (*Field Results Group*) en faveur d'un suivi et d'une gestion plus efficace des résultats. Il existe en particulier un besoin prouvé de :

- a) Développer les capacités, en particulier celle des partenaires de mise en œuvre de l'UNICEF et de leur personnel, pour concevoir, mettre en place et utiliser des systèmes de suivi et de gestion des résultats ;
- b) Améliorer la qualité de la conception des projets et des programmes, en particulier en ce qui concerne la mise au point de théories du changement claires et réalistes pour les programmes (voir la première recommandation) ;
- c) Établir des cibles utiles pour les résultats de programme et assurer le suivi de la performance des programmes au regard de ces cibles en utilisant les dispositifs mis en place le plus tôt possible en début du cycle de programme ;
- d) Fournir une assistance technique renforcée pour soutenir la collecte des données, le suivi et le rapportage des résultats par les partenaires.

**8. Concevoir et promouvoir une définition plus spécifique de l'innovation dans les programmes de l'UNICEF, couvrant à la fois les innovations programmatiques et technologiques.** Les évaluations examinées indiquent que les programmes soutenus par l'UNICEF ont souvent l'occasion de promouvoir des méthodes et des démarches innovantes permettant de renforcer la conception et la mise en œuvre

des programmes, le soutien technique et les mécanismes de prestation des services, ainsi que le contrôle et le suivi en continu. Dans de nombreux cas, ces innovations peuvent se révéler aussi importantes, voire plus importantes, que la promotion de l'utilisation de technologies avancées. Il existe également des possibilités pour renforcer la documentation de soutien à l'innovation de façon à ce que les décideurs nationaux disposent d'une meilleure information au moment de décider de financer ou non les innovations soutenues par l'UNICEF à l'échelle nationale. L'amélioration de la modélisation, de la collecte de données factuelles et du rapportage de l'UNICEF en matière de pratiques innovantes est également susceptible de renforcer la durabilité.

Prises ensemble, ces mesures devraient aider à assurer une amélioration continue des constatations obtenus et des conclusions documentés par la fonction d'évaluation de l'UNICEF pendant la durée du Plan stratégique actuel et de ceux qui lui succéderont.

# RESUMEN EJECUTIVO

## Antecedentes

En el presente informe se exponen los resultados del segundo estudio de la efectividad de los programas humanitarios y de desarrollo apoyados por el Fondo de las Naciones Unidas para la Infancia (UNICEF). El primer estudio, que se llevó a cabo en 2012 y 2013, analizó los resultados de las evaluaciones del UNICEF publicadas entre 2009 y 2011. El enfoque y las metodologías comunes a ambos estudios (2012 y 2016) se desarrollaron en el marco del documento Guía de la Red sobre Evaluación y Desarrollo del Comité de Asistencia para el Desarrollo (CAD-OCDE). Este Estudio se basa en el contenido de los informes de evaluación publicados por el UNICEF entre enero de 2012 y diciembre de 2015, así como en un análisis de documentos institucionales y de las consultas celebradas con el personal del UNICEF en su sede de Nueva York y en sus oficinas regionales (anexo F).

## Propósito

La finalidad de este estudio de la efectividad desde el punto de vista del desarrollo es generar un corpus de información creíble basado en un conjunto común de criterios (como se establece en el Documento Guía\*) que permita

conocer la efectividad de la labor humanitaria y de promoción del desarrollo del UNICEF. El estudio conllevó la realización de una síntesis sistemática de información a partir de una muestra de las evaluaciones disponibles publicadas entre 2012 y 2015 en lo que se refiere a los criterios de efectividad del desarrollo. El estudio también se propone aportar información para la toma de decisiones relativas, por ejemplo, a la introducción de ajustes en los sistemas de gestión basada en resultados, así como determinar los criterios de efectividad que requieren ajustes en la programación con el fin de mejorar los resultados.

## Enfoque y metodología

El estudio se llevó a cabo entre enero y agosto de 2016. Comenzó con una revisión preliminar de documentos del UNICEF y con la identificación de informes de evaluación del UNICEF existentes en la base de datos global de la organización, así como en su Sistema Global de Supervisión de Evaluación. También se realizaron entrevistas a personal clave en la sede del UNICEF en Nueva York y al personal de cada una de las oficinas regionales de la organización (anexo F).

En consulta con el personal de la Oficina de Evaluación, el equipo encargado del estudio identificó un universo potencial de 263 informes de evaluación. De este universo, se extrajo una muestra aleatoria modificada de 73 evaluaciones que, a continuación, el equipo encargado del estudio sometió a su propio control de calidad utilizando una plantilla de evaluación de la calidad de las evaluaciones, desarrollada específicamente para su uso en estudios de la efectividad del desarrollo (anexo D). De esos 73 informes, siete (un 9,6%) se eliminaron debido a las bajas puntuaciones que obtuvieron en este control de calidad, inferiores a la mínima exigida (25 puntos sobre 40).

### Siete esferas de resultados

1. Salud
2. VIH/SIDA
3. Agua, saneamiento e higiene
4. Nutrición
5. Educación
6. Protección infantil
7. Inclusión social

*Plan Estratégico 2014-2017 del UNICEF*

\* Assessing the Development Effectiveness of Multilateral Organizations: Guidance on the Methodological Approach (<http://www.oecd.org/dac/evaluation/dcdndep/50540172.pdf>)

Como resultado de ello, el examen se centró en 66 informes de evaluación, que se sometieron al correspondiente análisis sistemático, calificación y codificación de los hallazgos de las evaluaciones. Durante el proceso que se llevó a cabo en 2012, se analizaron 62 informes de evaluación.

Se examinó cada uno de los informes para evaluar sus conclusiones tomando como referencia seis criterios y 20 subcriterios de efectividad del desarrollo (anexo A). Las conclusiones obtenidas para cada subcriterio se clasificaron utilizando una escala de cuatro puntos, desde “altamente satisfactorio” (en el lado positivo) hasta “altamente insatisfactorio” (en el negativo). Durante el estudio se identificaron asimismo los factores que contribuyen tanto a los hallazgos positivos como a los negativos para cada uno de los seis criterios utilizados para evaluar la efectividad, según se reflejaban en los informes de evaluación. En los capítulos 3 y 4 de este informe se presentan los resultados detallados de este ejercicio de metasíntesis de las conclusiones de los informes de evaluación del UNICEF.

### Criterios de evaluación

1. Relevancia de la intervención
2. Logro de los objetivos humanitarios y de desarrollo
3. Temas transversales (promoción de la igualdad de género y la equidad)
4. Sostenibilidad de los resultados y beneficios
5. Eficiencia
6. Utilización de la evaluación y el seguimiento para mejorar la efectividad

### Cobertura de los criterios de evaluación en las evaluaciones examinadas

El estudio establecía una serie de categorías para evaluar el grado de cobertura de los 20 subcriterios en las 66 evaluaciones incluidas en la muestra, con base en el número de evaluaciones en las que se abordaba adecuadamente cada uno de los subcriterios. La calificación

de la cobertura podía ser “elevada”, “moderada” o “escasa”. Las evaluaciones examinadas durante el ejercicio llevado a cabo en 2016 lograron un nivel de cobertura notablemente superior al de las examinadas en 2012.

### Calificaciones de la cobertura

- Elevada:** los subcriterios estaban recogidos en 46 informes de evaluación o más
- Moderada:** los subcriterios estaban recogidos entre 31 y 45 informes de evaluación
- Escasa:** los subcriterios estaban recogidos en menos de 31 evaluaciones

### Limitaciones

Las principales limitaciones del estudio están relacionadas con el muestreo y con la naturaleza retrospectiva de la evaluación y de la metasíntesis en las cuales se basa la metodología.

*Sesgo de muestreo:* pese a que la muestra de evaluaciones comenzó con un muestreo aleatorio estratificado, se ajustó intencionalmente con el fin de mejorar la representación de las evaluaciones pertenecientes al sector de la salud y de las procedentes de países del África subsahariana. Esto dio lugar a una muestra intencional que pretendía ser ilustrativa de los resultados de las evaluaciones del UNICEF publicadas entre 2012 y 2015. No se trata, por tanto, de una muestra aleatoria válida de evaluaciones del UNICEF desde el punto de vista estadístico que se pueda utilizar para calcular intervalos de confianza.

*Limitaciones de la estratificación:* la muestra de evaluaciones se equilibró con el fin de que reflejara de manera amplia la distribución de los informes de evaluación elaborados por el UNICEF por área geográfica, año de publicación y esfera de resultados del plan estratégico. No obstante, los estratos derivados de la muestra son excesivamente reducidos como para obtener resultados significativos dentro de cada subcomponente (para las diferentes esferas de resultados, por ejemplo).

*Evaluación de la eficacia de los programas humanitarios.* Las limitaciones señaladas referidas a la estratificación de la muestra, adquieren una importancia especial a la hora de analizar la dimensión humanitaria del trabajo del UNICEF. Dado que los informes de evaluación con un enfoque centrado directamente en los programas humanitarios representaban tan solo un 2,7% del total de evaluaciones publicadas del UNICEF durante el período objeto del estudio, solamente se pudo incluir un número muy reducido de ellos en el estudio si se quería que la muestra reflejara la población global. Si bien las evaluaciones centradas directamente en el aspecto humanitario recibieron un peso relativamente desproporcionado en la muestra (un 6,1% de las evaluaciones incluidas en ella), su número sigue siendo demasiado limitado como para permitir cualquier tipo de generalización de los resultados a la experiencia global del UNICEF en el ámbito de la programación humanitaria. Es importante señalar que el universo de evaluaciones, y también la muestra, incluyen evaluaciones clasificadas dentro de otras esferas de resultados (como agua y saneamiento o programas sanitarios), que tienen lugar en contextos humanitarios o de emergencia y, por tanto, pueden tener consecuencias para los programas humanitarios. No obstante, el proceso de revisión de las evaluaciones incluidas en la muestra no aportó pruebas suficientes sobre las operaciones humanitarias como para respaldar la formulación de conclusiones en el ámbito de la programación humanitaria.

*Naturaleza retrospectiva de la metátesis:* por definición, los informes de evaluación son de carácter retrospectivo, en el sentido de que proporcionan evaluaciones ex ante de la efectividad de los programas. Muchos de los programas evaluados entre 2012 y 2015 comenzaron muchos años antes de ser evaluados. Esto significa que el estudio debe tener cuidado de no responsabilizar a los programas más antiguos de las nuevas estrategias, o de extrapolar los resultados históricos a la práctica actual.

Pese a que las limitaciones expuestas son reales, ninguna de ellas socava en exceso la utilidad de los resultados del estudio de la

efectividad del desarrollo. Al prestar atención a la selección de la muestra de evaluaciones y a la clasificación y análisis de los resultados de estas, el equipo encargado del estudio tuvo la posibilidad de reproducir el proceso que se siguió en 2012 y de proporcionar un visión significativa de la evidencia evaluativa existente con respecto de la efectividad organizacional del UNICEF (y cómo ha cambiado esta en el período de ejecución de las intervenciones).

### Iniciativas recientes del UNICEF

Las conclusiones y recomendaciones que se exponen en el presente informe no deben interpretarse como una indicación de que el UNICEF desconociera o no tratara de dar respuesta a los hallazgos de los informes de evaluación publicados entre 2012 y 2015. Las iniciativas recientes y en curso de la organización de las que se ha informado al equipo encargado de la realización del estudio, lideradas por el Grupo de Resultados sobre el Terreno y que buscan fortalecer la programación (incluida la gestión basada en resultados) incluyen:

- Iniciativas de capacitación del UNICEF en materia de gestión basada en resultados a escala global (2015-2017), en las que en 2016 participaron 1.500 trabajadores de la organización y se prevé que participen un total de 4.000 hasta finales de 2017. Para estas iniciativas se utiliza un equipo global de capacitadores y un conjunto de módulos de capacitación en persona y en modalidad de aprendizaje electrónico;
- Notas sobre la estrategia de los programas: se introdujeron en noviembre de 2015 y se incluyeron en todos los nuevos programas de país. El personal informa de que se han desarrollado más de 85 de estas notas, cada una de las cuales incluye una teoría de cambio, un marco de resultados y la adecuación de los resultados y estrategias a la financiación prevista. En la actualidad se está realizando un examen de la calidad de las notas sobre estrategia elaboradas en 2016;

- La aplicación del cuadro de control del desempeño de las oficinas de país con un vínculo directo entre los resultados de estas oficinas, medidos a través de dicho cuadro de control, y la evaluación de personal del desempeño del Representante del UNICEF.

No obstante, los términos de referencia del estudio y la metodología empleada no incluían ni permitían ninguna iniciativa dirigida a evaluar la efectividad de las medidas anteriores a la hora de abordar las cuestiones que aquí se reflejan.

## Conclusiones

Los hallazgos detallados del estudio realizado en 2016 sobre la efectividad del desarrollo del UNICEF se presentan en los capítulos 3 y 4 del informe. En esta sección se exponen las conclusiones del estudio.

### Relevancia

1. Al igual que en 2012, el estudio de la efectividad del desarrollo de las evaluaciones del UNICEF **demuestra que los programas del UNICEF siguen caracterizándose por ser altamente relevantes**. En particular, las evaluaciones realizadas entre 2012 y 2015 señalaban que el UNICEF respaldó **proyectos y programas correctamente alineados con las prioridades y programas de desarrollo nacional**. Si bien las evaluaciones reportan hallazgos positivos en términos de relevancia, también identifican **mejoras potenciales** en cuanto a la relevancia de algunos programas a través de la **implicación de las partes interesadas clave durante el desarrollo de los programas así como a través de vínculos más sólidos entre evaluación de necesidades y la focalización de la asistencia**.

### Logro de objetivos

2. **Los programas apoyados por el UNICEF que se evaluaron entre 2012 y 2015 lograron los objetivos que se habían fijado**. Las

evaluaciones examinadas demuestran que los programas respaldados por el UNICEF alcanzaron un nivel de logro de objetivos muy satisfactorio y moderadamente superior al constatado en el estudio realizado en 2012. En particular, **en muchos informes de evaluación se señalaba que el UNICEF había logrado resultados positivos en su trabajo a nivel macro orientado al apoyo** del desarrollo de políticas y estrategias; y en el apoyo a programas eficaces de capacitación y desarrollo de capacidades. Estos aspectos habían sido señalados como **ámbitos de mejora en el estudio de 2012**: el anterior estudio sobre la efectividad del desarrollo identificó como un factor negativo, el limitado trabajo en materia de desarrollo de capacidades y lo señaló como un factor negativo que en ocasiones limitaba la consecución de los objetivos. Por otro lado, las evaluaciones examinadas señalaban asimismo que algunos programas no habían alcanzado los objetivos planteados como consecuencia de **debilidades en su diseño, sobre todo a la ausencia de teorías de cambio claras y correctamente formuladas**.

### Promoción de la igualdad de género

3. Las **evaluaciones publicadas por el UNICEF demostraron una probabilidad muy superior de abordar las cuestiones relativas a la igualdad de género** en el período objeto de estudio (2012-2015) que los programas ejecutados entre 2009 y 2011, según queda reflejado en el estudio de 2012. La mayor atención prestada a la igualdad de género en las evaluaciones publicadas entre 2012 y 2015 sugiere que la **función de evaluación del UNICEF ha respondido positivamente al Plan de Acción sobre Prioridades Estratégicas** en materia de igualdad de género (2010 a 2013), así como a los resultados obtenidos por el estudio de la efectividad del desarrollo llevado a cabo en 2012.
4. Aunque la cobertura en cuanto al número de informes de evaluación publicados por el UNICEF ha mejorado desde el estudio

de 2012, **los hallazgos referentes a la efectividad en el abordaje de la igualdad de género, siguen siendo muy decepcionantes, según lo expuesto en dichos informes**. El porcentaje de evaluaciones analizadas en 2016 que reportan hallazgos positivos en el ámbito de la igualdad de género (48%) es prácticamente idéntico al expuesto en el estudio de 2012 (47%). Ahora, como entonces, **menos de la mitad de los informes de evaluación de programas del UNICEF que abordan la igualdad de género como tema transversal reportan hallazgos positivos** en este terreno. Esta aparente falta de mejora resulta sorprendente en el marco del Plan de Acción sobre Prioridades Estratégicas 2010-2013.

5. **La ausencia de una perspectiva de género o la falta de aplicación de un análisis de género durante el diseño y la ejecución de los programas** es de lejos, el motivo más común, señalado en los informes de evaluación, que explica la ausencia de un abordaje adecuado de la igualdad de género como tema transversal. Este fue también el factor citado con mayor frecuencia como limitante de la efectividad a la hora de abordar la igualdad de género, en el estudio de 2012. **Estas conclusiones plantean preguntas muy importantes acerca del adecuado equilibrio entre la transversalización de la igualdad de género y las necesidades de intervenciones selectivas específicas dirigidas a abordar la cuestión del género en los programas apoyados por el UNICEF.**

#### **El enfoque del UNICEF con respecto a la equidad**

6. Las evaluaciones publicadas por el UNICEF entre 2012 y 2015 ofrecen poderosos indicios de que **la función de evaluación ha respondido adecuadamente a la reorientación del enfoque en lo que respecta a la equidad**, que comenzó en 2010. Una significativa mayoría de las evaluaciones examinadas (54 de 66) abordaba directamente la cuestión de si los

programas del UNICEF promueven eficazmente la mejora de la equidad en beneficio de los niños y adolescentes.

7. Los resultados de este nuevo subcriterio relativo a la **efectividad para la mejora de la equidad, fueron ampliamente positivos**. Los programas respaldados por el UNICEF consiguieron abordar la equidad al dirigirse a un conjunto diverso de grupos vulnerables o desfavorecidos, entre los que figuraban los hogares situados en los estratos socioeconómicos más bajos, los distritos rurales más distantes y a los que resulta más complicado llegar, los grupos étnicos o tribales desfavorecidos, los trabajadores migrantes, los niños trabajadores en entornos urbanos y las adolescentes que se enfrentan a barreras específicas para asistir a la escuela. También parece evidente, no obstante, que **para abordar eficazmente la equidad es necesario que tanto el UNICEF como sus socios identifiquen con claridad a las personas desfavorecidas como grupo destinatario, ubiquen correctamente a este grupo y detecten los obstáculos que le impiden acceder a los servicios o recursos, con el fin de mejorar la adecuación de las intervenciones.**
8. Continúa existiendo un margen considerable para mejorar la efectividad de una proporción significativa de los programas apoyados por UNICEF, en materia de desigualdad. En la mayor parte de los casos, esta mejora requeriría que los programas apoyados por el UNICEF **desarrollen una estrategia explícita para centrarse en beneficiar a las personas y comunidades marginadas; una estrategia que vaya acompañada de criterios de participación y de acceso a los servicios que proporcione una ventaja clara a esas personas y comunidades**. Esta estrategia basada en dirigir las intervenciones hacia los grupos marginados sería necesaria para el trabajo normativo preliminar así como para los programas destinados a los beneficiarios.

## Sostenibilidad

9. Al igual que con los temas de igualdad de género, **la función de evaluación del UNICEF ha mejorado el nivel de atención prestada a las cuestiones relativas a la sostenibilidad durante el período comprendido entre 2012 y 2015** en comparación con lo ocurrido con las evaluaciones realizadas de 2009 a 2011.
10. Pese a la mejora del nivel de cobertura, las evaluaciones examinadas en 2016 siguen reflejando hallazgos **poco satisfactorios en lo referente a la sostenibilidad**. En comparación con los resultados de 2012, se observan algunas mejoras en los hallazgos relativos al fortalecimiento de la capacidad institucional y comunitaria, así como en lo que concierne al fortalecimiento del entorno normativo. No obstante, la pregunta básica sobre si los beneficios de los programas del UNICEF se mantendrán o no tras la finalización de los programas continúa obteniendo respuestas negativas en muchos informes de evaluación (concretamente en un 74% de ellos). Parece claro que, para que los proyectos y programas apoyados por el UNICEF eleven la sostenibilidad de sus resultados, será necesario ir más allá de las mejoras señaladas en términos de desarrollo de las capacidades y promoción de un entorno favorable. **Los informes de evaluación apuntan directamente a la necesidad de que los proyectos y programas respondan eficazmente a la necesidad de planificar y garantizar la sostenibilidad financiera y programática.**

## Eficiencia

11. La mayor parte de los programas del UNICEF evaluados desde 2012 resultaron ser **eficientes en cuanto al control de los costos y la utilización de los recursos**. Estos programas también **se ejecutaron y lograron sus objetivos dentro de los plazos previstos**. Sin embargo, como ocurrió en el caso del estudio de 2012, las evaluaciones indican la necesidad de prestar una atención constante a la eficiencia de los programas respaldados

por el UNICEF en términos de costos y utilización de recursos. Asimismo, hacen hincapié en la necesidad de reforzar los sistemas y procesos de seguimiento de los costos y resultados de los programas, con el fin de gestionarlos de forma eficiente.

## Evaluación para mejorar la efectividad

12. El estudio tomó nota de la existencia de constantes **mejoras en el uso de la evaluación, por parte del UNICEF, como función que puede contribuir a incrementar la efectividad de los programas**. Estas incluyen, mejoras continuas en la calidad y utilidad de los informes de evaluación, como queda reflejado en las respuestas proporcionadas por los miembros de la dirección. Pese a ello, sigue siendo necesario **garantizar una cobertura mayor y un enfoque sistemático con respecto al establecimiento de prioridades de los programas y proyectos que se evaluarán, de modo que los programas más importantes se sometan a evaluaciones periódicas y de alta calidad.**

## Seguimiento y gestión basada en resultados para mejorar la efectividad

13. Si bien el UNICEF ha invertido **esfuerzos sustanciales para mejorar la solidez de sus sistemas de seguimiento y el uso de métodos de gestión basada en resultados** desde 2012, esto **no se ha traducido en mejoras apreciables** en la fortaleza de los sistemas de seguimiento según reportan las evaluaciones llevadas a cabo entre 2012 y 2015. Esto puede deberse a que el Grupo de Resultados sobre el Terreno se constituyó hace relativamente poco tiempo (2014), pero también indica que el fortalecimiento de los sistemas de seguimiento y gestión basada en resultados continúa representando un desafío para el UNICEF.
14. La mejora de la gestión, el seguimiento de los resultados y la elaboración de informes por parte del UNICEF requiere la **adopción de una serie de medidas** por parte de este

organismo, como el **desarrollo de capacidades** de su personal y, especialmente, de sus **asociados en la ejecución**; la mejora de la claridad de las **teorías de cambio** en las que se sustenta el diseño de los programas; el establecimiento de **metas realizables en los programas**, que sean objeto de **seguimiento a través de sistemas implementados lo antes posible en el ciclo programático**; y la prestación de **asistencia técnica** para respaldar la recogida de datos y el seguimiento de los resultados por parte de los asociados en la ejecución. El mandato del Grupo de Resultados sobre el Terreno responde directamente a muchos de estos desafíos.

### Innovación

15. Las evaluaciones examinadas indican que el UNICEF ha realizado **progresos en el impulso de un compromiso con la innovación**, que incluye la disposición a experimentar con diferentes modelos de apoyo a programas y prestación de servicios, así como a utilizar tecnologías emergentes de formas innovadoras. Sin embargo, también se aprecian indicios de que el **énfasis en la innovación se puede fortalecer a través del desarrollo de una definición clara de lo que se entiende por “programas innovadores” y prestando mayor atención a las innovaciones en la fase de diseño de los programas y servicios**. También existe margen para incrementar el peso otorgado a la innovación, a través de la documentación y la aplicación a escala de aquellas innovaciones que hayan demostrado ser efectivas.

### Enfoques basado en los derechos humanos en la programación

16. Las evaluaciones del UNICEF publicadas entre 2012 y 2015 demuestran **un mayor interés en los enfoques basados en los derechos humanos en la programación**. Aunque dichas evaluaciones indican que numerosos programas han realizado un intento serio de adoptar un enfoque basado en derechos humanos, también señalan que algunos de

ellos muestran un compromiso más retórico que real. En particular, los informes de las evaluaciones apuntaban a **la necesidad de abordar tanto las capacidades de los titulares de obligaciones para prestar servicios como la capacidad de los titulares de derechos para demandar eficazmente dichos servicios**.

En resumen, los hallazgos de las evaluaciones publicadas por el UNICEF en el período 2012-2015 evidencian mejoras en cuanto a la efectividad global de estos programas desde el punto de vista del desarrollo, así como en el abordaje de cuestiones clave, a través de la evaluación. Al mismo tiempo, las evaluaciones examinadas señalan una serie de esferas susceptibles de mejora en aras de la efectividad, como, por ejemplo, la programación en favor de la igualdad de género y de la sostenibilidad.

### Recomendaciones

Las recomendaciones que siguen están basadas en las conclusiones expuestas en el informe e indican que el UNICEF debería:

1. **Fortalecer la capacidad del UNICEF y de sus socios para mejorar el diseño de sus programas apoyándose en una teoría de cambio sólida**. Pese a que los programas apoyados por el UNICEF suelen alcanzar sus objetivos, el estudio también identifica áreas que merecen mejorar, especialmente el desarrollo de teorías de cambio válidas que especifiquen adecuadamente los vínculos entre las actividades y los recursos de los programas, por un lado, y los productos y resultados de estos, por otro. Esto debería incluir:
  - a) iniciativas de desarrollo de capacidades dirigidas al personal del UNICEF y de los socios en las oficinas de país en temas relativos al desarrollo de teorías de cambio eficaces y robustas desde el punto de vista técnico durante la fase de diseño de los programas; estas iniciativas deberían incluir la identificación de las hipótesis causales clave que podrán

verificarse durante las evaluaciones. Esto requerirá una orientación concreta y mejorada sobre la elaboración de teorías de cambio eficaces;

- b) desarrollo de capacidades y asistencia técnica constante al personal del UNICEF y a los socios en la identificación de metas realizables para los programas y de indicadores apropiados para el seguimiento del logro de dichas metas a lo largo del tiempo;
  - c) la vinculación del trabajo para mejorar del diseño de los programas con iniciativas dirigidas a reforzar los sistemas de seguimiento y gestión basados en resultados (véase la recomendación núm. 7).
2. **Aplicar de manera más coherente los métodos de análisis de género e incorporar la perspectiva de género en el diseño de los programas y en todas las fases del ciclo de programación.** La ausencia de mejoría en los resultados del estudio de 2016 en materia de igualdad de género en comparación con el estudio de 2012 indica la necesidad de:
- a) garantizar que los procesos de diseño y los sistemas de aprobación de los programas incorporen plenamente el análisis de género y la perspectiva de género antes de que el UNICEF asuma compromiso alguno con dichos programas;
  - b) fortalecer la capacidad del personal del UNICEF y sus socios para utilizar herramientas de análisis de género durante el diseño de los programas;
  - c) mejorar la recogida y utilización de datos desglosados por género en los programas del UNICEF;
  - d) consultar a las comunidades sobre la situación de las mujeres y las niñas durante el diseño de los programas e incorporar, durante los procesos de consulta, a organizaciones que representen los intereses de mujeres y niñas;

e) equilibrar los enfoques orientados a la transversalización de género con intervenciones selectivas dirigidas específicamente a hacer frente a las desigualdades de género (incluida la incorporación de objetivos específicos, claros y pertinentes desde el punto de vista del género durante la fase de diseño de los programas); y

f) seguir prestando mayor atención a la igualdad de género como tema transversal en los sistemas y procesos liderados por el UNICEF para el seguimiento y la evaluación, con el fin de continuar avanzando a partir de los progresos reportados entre los estudios de 2012 y 2016.

Estas medidas son congruentes con el enfoque más selectivo adoptado por el UNICEF en su Plan de acción en materia de género de 2014, sobre todo si se tiene en cuenta el peso otorgado a la necesidad de que los programas identifiquen y traten de eliminar los obstáculos y cuellos de botella existentes.

3. **Definir e identificar más claramente los grupos de mujeres y niños desfavorecidos destinatarios de los programas apoyados por el UNICEF** y aclarar los beneficios específicos que el programa ofrecerá a este grupo en comparación con otros colectivos. Los resultados del estudio indican que existe una oportunidad para seguir progresando a partir del actual nivel de resultados positivos, a través del perfeccionamiento del enfoque de equidad. Esto debería incluir:
- a) esfuerzos más constantes para identificar a los grupos más desfavorecidos de mujeres y niños durante la fase de diseño de los programas, incluida una detección más precisa de su ubicación en las comunidades atendidas;
  - b) una mejor identificación de las barreras a la participación y acceso a los servicios a las que se enfrentan los grupos de mujeres y niños desfavorecidos definidos

como colectivo destinatario, de modo que los programas vayan expresamente dirigidos a ellos;

- c) criterios selectivos para la participación en los programas y el acceso a los servicios para los grupos desfavorecidos identificados, incluidos criterios que prioricen el acceso de estos grupos y anticipen específicamente un resultado desproporcionadamente positivo en su favor;
  - d) mejora del seguimiento de la participación, el acceso y los beneficios generados en favor de los miembros de los grupos desfavorecidos de mujeres y niños definidos como colectivos destinatarios de los programas. Esto puede incluir la garantía de que los programas seleccionados respondan a la necesidad de reforzar los sistemas nacionales de gestión de los datos para mejorar los resultados en términos de equidad.
4. **Abordar de un modo más coherente y eficaz la sostenibilidad financiera y programática en el diseño de los programas y estrategias de los diferentes países.** Pese a que se ha producido un avance notable en términos de sostenibilidad institucional de los programas respaldados por el UNICEF, existe considerable evidencia de que la sostenibilidad financiera de las iniciativas apoyadas representa un desafío significativo para la sostenibilidad de los proyectos y programas impulsados por el UNICEF. El UNICEF debería hacer todo lo que esté en su mano, tanto durante la fase de diseño de los programas como durante su ejecución, para:
- a) desarrollar y aplicar mecanismos generadores de ingresos para los proyectos apoyados, de modo que estos puedan mantenerse una vez que finalice la ayuda del UNICEF;
  - b) cuando no resulte apropiado establecer ese tipo de mecanismos, fomentar que los gobiernos nacionales, provinciales y locales de los países en los que

se ejecuten los proyectos y programas doten las partidas presupuestarias necesarias; y

- c) abogar por y contribuir a obtener fuentes de financiación alternativas a través de organismos multilaterales, donantes bilaterales o el sector privado.
5. **Continuar con (o incluso intensificar) los esfuerzos dirigidos a mejorar la eficiencia de los programas respaldados por el UNICEF.** A medida que un mayor número de evaluaciones aborda las cuestiones relativas al costo y la eficiencia en los programas apoyados por el UNICEF, resulta evidente que existe margen para mejorar la adecuación y el nivel de atención prestado a los temas de la eficiencia y la eficacia en función de los costos en la fase de diseño de los programas. El requisito más urgente es reforzar el diseño de los programas y los sistemas de seguimiento de estos de manera que permitan supervisar los costos de los programas a lo largo del tiempo y establecer un vínculo creíble entre los costos unitarios y los resultados. De lo contrario no cabe esperar que los gestores de los programas sean capaces de gestionarlos en aras de lograr una mayor eficiencia. Los esfuerzos dirigidos a mejorar la eficiencia también podrían beneficiarse del desarrollo de una metodología básica común para evaluar la eficiencia de los programas en el UNICEF.
6. **Mejorar la selección de los programas para su evaluación en el UNICEF con el objetivo de garantizar mayores niveles de cobertura de los programas e introducir mecanismos más sistemáticos de adopción de decisiones en lo referente a las prioridades.** Pese a que la Política de evaluación aprobada recientemente recomienda evaluar todos los programas al menos una vez cada cinco años, existe margen para establecer una vinculación más precisa entre la decisión relativa a qué programas y proyectos deberían evaluarse primero (dados los limitados recursos disponibles) y los requisitos

asociados tanto al Plan Estratégico general del organismo como a la necesidad de mejorar la efectividad de los programas. Esto es especialmente importante en las oficinas de país, donde la lógica de selección de programas y proyectos para su evaluación no siempre está clara. También puede conducir a prestar una mayor atención a los nexos existentes entre los programas humanitarios y de desarrollo apoyados por el UNICEF. El mecanismo del Plan de Evaluaciones con Costos (PEC) puede representar un canal muy importante para fortalecer el proceso de priorización de las decisiones de evaluación en las oficinas de país.

**7. Proseguir con (e intensificar) los esfuerzos dirigidos a fortalecer los sistemas de gestión, seguimiento de los resultados y presentación de informes.** Esto pone de relieve la naturaleza crucial del mandato del Grupo de Resultados sobre el Terreno en apoyo a una gestión y un seguimiento más eficaces de los resultados. En particular, se ha demostrado que es necesario:

- a) desarrollar la capacidad, en particular, de los asociados del UNICEF en la ejecución y de su personal en materia de diseño, establecimiento y utilización de sistemas de gestión y seguimiento de los resultados;
- b) mejorar la calidad básica del diseño de los programas y proyectos, sobre todo en lo que se refiere al desarrollo de teorías de cambio claras y realistas para los programas (véase la recomendación núm. 1);
- c) establecer metas realistas en lo que se refiere a los resultados de los programas y llevar a cabo un seguimiento de dichos resultados tomando como referencia las metas establecidas, utilizando para ello sistemas implantados en las fases más tempranas posibles del ciclo de programación; y

d) prestar una asistencia técnica de mayor calidad para respaldar la recogida de datos, el seguimiento de los resultados y la presentación de informes por parte de los asociados en la ejecución.

**8. Desarrollar y promover una definición más específica de innovación en los programas del UNICEF, que abarque tanto las innovaciones programáticas como las tecnológicas.** Las evaluaciones examinadas indican que los programas apoyados por el UNICEF se encuentran a menudo con oportunidades para promover la adopción de enfoques y métodos innovadores para mejorar el diseño y la ejecución de los programas, el apoyo técnico y la prestación de servicios, así como el seguimiento y la supervisión permanentes. En muchos casos, esto puede ser tan o más importante que promover el uso de tecnologías avanzadas. También existen oportunidades para perfeccionar la documentación de apoyo a la innovación, de manera que los responsables de la toma de decisiones a escala nacional dispongan de información de mayor calidad a la hora de decidir sobre la pertinencia (o no) de financiar innovaciones respaldadas por el UNICEF en el ámbito nacional. La mejora de los modelos, la recopilación de evidencia y la presentación de informes sobre prácticas innovadoras por parte del UNICEF también podría redundar positivamente en la sostenibilidad.

En conjunto, estas medidas deberían contribuir a garantizar la mejora continua de los hallazgos y conclusiones reportados por la función de evaluación del UNICEF a lo largo de la vigencia de los actuales y eficaces Planes Estratégicos del organismo.



CHAPTER 1  
**INTRODUCTION**

# CHAPTER 1

## 1.1 BACKGROUND

There has been increased attention to development effectiveness not only just at the United Nations Children's Fund (UNICEF), but also in all United Nations (UN) agencies, funds and programs among donors. The most recent Quadrennial Comprehensive Policy Review (QCPR)<sup>2</sup> of operational activities for the development of the UN systems, including UNICEF, encompasses a pronounced focus on development effectiveness<sup>3</sup>. UNICEF's Strategic Plan 2014-2017 also highlights the need to continue accelerating work in this area.

In order to continue progress in learning more about the development effectiveness of UNICEF, the Evaluation Office has commissioned a review to synthesize findings from past evaluations. The findings and lessons from this review are expected inform on-going policy, strategy and program development in order to best achieve results for children. The first development effectiveness review (DER) of UNICEF was completed in 2013 by the Government of Netherlands. It was carried out using a methodology and approach endorsed by the OECD DAC Network on Development Evaluation and covered the period 2009-2011. As the bulk of the work for that study was carried out during 2012, it is referred to in this report as the 2012 review.

This report presents the results of the second review of the development and humanitarian effectiveness of UNICEF. The review relies on the content evaluation reports produced by UNICEF and published between January 2012 and December 2015. The information gathered from these evaluations was supplemented by a review of corporate documents

and consultations with staff at UNICEF headquarters in New York and in regional and country offices (Annex F).

## 1.2 PURPOSE

The purpose of this DER is to generate a body of credible information on a common set of criteria (as prescribed in the Guidance Document<sup>4</sup>) that provides an insight into the development and humanitarian effectiveness of UNICEF. It involved conducting a systematic synthesis of information from a sample of available evaluations published from 2012 to 2015 as they relate to the criteria of development effectiveness. The review is also intended to inform decisions on issues such as adjustments to results-based management (RBM) systems and determining those effectiveness criteria which require adjustments in programming for more positive results.

The DER report is envisaged to be used in the following ways:<sup>5</sup>

- By UNICEF managers, to gain an independent external view of UNICEF's development effectiveness and consequently inform decisions;
- By members of the UNICEF Executive Board and UNICEF partners, especially those requiring evidence of how well UNICEF is performing. This review is also one of the commitments made to and accepted by the Executive Board in the global plan of thematic evaluations discussed at the 2/2014 Board session; and,
- By stakeholders in UNICEF's Evaluation Office.

<sup>2</sup> QCPR (2012) [http://www.un.org/ga/search/view\\_doc.asp?symbol=E/2012/79](http://www.un.org/ga/search/view_doc.asp?symbol=E/2012/79)

<sup>3</sup> The General Assembly adopted on 21 December 2012 a landmark resolution (67/226) on the quadrennial comprehensive policy review (QCPR) of UN operational activities for development.

<sup>4</sup> Assessing the Development Effectiveness of Multilateral Organizations: Guidance on the Methodological Approach (<http://www.oecd.org/dac/evaluation/dcdndep/50540172.pdf>)

<sup>5</sup> Statement of Work and Terms of Reference for the Development Effectiveness of UNICEF, December 14, 2015

### 1.3 STRUCTURE OF THE REPORT

The report is structured as follows:

- **CHAPTER ONE** provides an introduction to the review and a general description of UNICEF's strategic direction, key output areas and overall operations;
- **CHAPTER TWO** presents a brief description of the approach and methodology used to carry out the review;
- **CHAPTER THREE** details the findings of the review in relation to the six main criteria and 20 sub-criteria used to assess effectiveness in development programming and humanitarian action. Each section of Chapter Three (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 allow the reader to understand the context for the findings which follow. Each short sub-section on coverage is then followed by a detailed discussion of the results found in the evaluations reviewed:
- **CHAPTER FOUR** focuses more specifically on the implications of the review for results monitoring and for evaluation at UNICEF; and,
- **CHAPTER FIVE** presents the findings of the review in relation to innovation and human rights-based approaches.
- **CHAPTER SIX** presents the conclusions and recommendations of the review.

### 1.4 STRATEGIC DIRECTION AND OPERATIONS AT UNICEF

#### Strategic Direction: The Rights of Every Child

The UNICEF Strategic Plan 2014-2017 focuses squarely on realizing the rights of every child,<sup>6</sup> especially the most disadvantaged. The ultimate

mission of UNICEF is to promote the rights of every child, everywhere, in everything the organization does – in programs, in advocacy and in operations. The equity strategy, emphasizing the most disadvantaged and excluded children and families, translates this commitment to children's rights into action. For UNICEF, equity means that all children have an opportunity to survive, develop and reach their full potential, without discrimination, bias or favouritism.<sup>7</sup>

The Strategic Plan has been informed by major global movements and discussions about the Sustainable Development Goals (SDGs) and the lessons learned and findings from the review of the Medium-Term Strategic Plan (MTSP) 2006-2013.<sup>8</sup>

The Strategic Plan enables UNICEF to contribute to and foster coherence and synergies among all funds, program and specialized agencies in the United Nations development system, to better support countries. It also enables UNICEF to leverage a growing diversity of partnerships with government, civil society, the private sector and others. These include public-private partnerships, South-South and triangular cooperation, and engagement with non-governmental organizations (NGOs).

The Strategic Plan bridges the final two years of the MDG timetable and the post 2015 Sustainable Development Goals (SDGs). It focuses the work of UNICEF on seven outcomes (see Box 1).

Humanitarian action, gender equality and human rights cut across all seven outcomes. These seven outcomes represent areas of institutional and behavioral change consistent with the Results-Based Management (RBM) handbook of the United Nations Development Group (UNDG). They are the core results to which UNICEF contributes in collaborations with

<sup>6</sup> UNICEF mandate covers all children below the age of 18 years, as the Convention on the Rights of the Child.

<sup>7</sup> The UNICEF Strategic Plan 2014-2017.

<sup>8</sup> The five focus areas of the MTSP were: young child survival and development, basic education and gender equality, HIV/AIDS and children, child protection, and policy and advocacy and partnerships for children's rights.

**BOX 1****Seven Outcome Areas**

1. Health
2. HIV&AIDS
3. Water Sanitation and Hygiene (WASH)
4. Nutrition
5. Education
6. Child Protection
7. Social Inclusion

*UNICEF Strategic Plan 2014-2017*

national partners, civil society organizations and other stakeholders. The seven outcomes of the Strategic Plan 2014-2017 are envisaged to encompass the major aspects of child well-being. The outcomes also encompass the combined results of development and humanitarian programming.

Each outcome has a common set of six outputs which are achieved by the work of UNICEF and its partners.<sup>9</sup>

**BOX 2****Common Set of Six Outputs for Each Outcome**

**Demand** - Enhanced support for children, families and communities to promote knowledge, behaviour change, demand for services and opportunities for participation;

**Supply** - Increased national capacity to ensure availability of and access to services and to strengthen systems;

**Enabling environment** - Strengthened political commitment and national capacity to legislate, plan and budget for children;

**Humanitarian action** - Increased country capacity and delivery of services to protect and provide for children in humanitarian settings;

**Rights and gender equality** - Increased capacity of governments and partners, as duty-bearers, to identify and respond to specific challenges related to the protection and promotion of the rights of children and gender equality;

**Global and regional action** - Enhanced enabling environment at global and regional levels to realize child rights.

**Operations, Revenues and Expenditures**

In 2015, UNICEF worked in more than 190 countries, territories and areas through country programs and National Committees organized through a network of seven regional offices (ROs) and 128 country offices (COs). UNICEF's 2014 gross revenues provided by major donors, governments, private foundations, corporate alliances and other intergovernmental and other partners amounted to USD 5,169 million. This included USD 1,579 million received towards humanitarian assistance in 2014 (an increase of 19 percent compared to 2013).<sup>10</sup>

The largest share of program expenses (USD 4,131 million) in 2014 was directed to programs in health (29.8 percent) followed by education (20.0 percent) and water, sanitation and hygiene (WASH) (17.6 percent) as illustrated in Figure 1. The resources for the program expenditures were categorized as regular resources (20.8 percent), other regular resources (50.1 percent; and other emergency resources (29.1 percent).

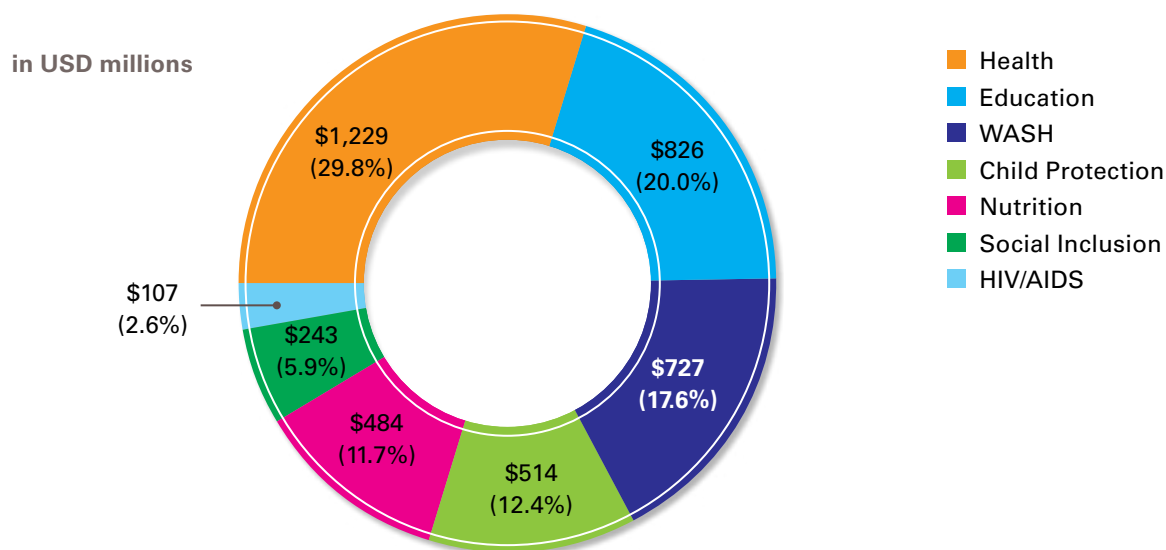
<sup>9</sup> UNICEF Strategic Plan 2014-2017.

<sup>10</sup> UNICEF Annual Report 2014.

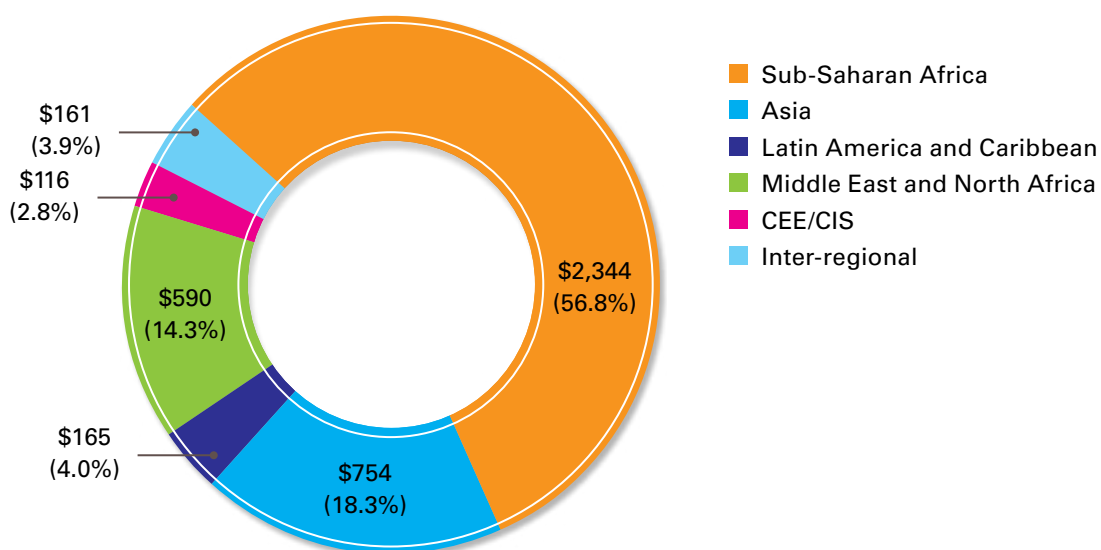
In terms of regional focus (Figure 2), a little more than half of the program expenses were directed towards sub-Saharan Africa, the region with the largest number of low-income countries. With more countries transitioning into lower or upper middle-income status, the proportion of program expenditure has marginally declined in Asia, Latin America and

Caribbean (LAC) and Central and Eastern Europe and Commonwealth of Independent States (CEE/CIS) regions. However, due to the emergency situation in Middle-East Asia and North Africa (MENA) region, expenditures there have increased from USD 143 million in 2012 to USD 590 million in 2014 (from 4.9 percent of 2012 program expenditure to 14.3 percent in 2014).

**FIGURE 1 Program Expenses by Output Area (USD 4,131 million) 2014**



**FIGURE 2 Program Expenses by Region (USD 4,131 million) 2014**



## 1.5 EVALUATION AT UNICEF

In recent years, the UNICEF Executive Board has paid particular attention to the evaluation function within UNICEF, continuing to affirm its central role. In decision 2013/3, the Executive Board approved the revised Evaluation Policy (E/ICEF/2013/14) and emphasized that the evaluation function would continue to report directly to the Executive Board.

## 1.6 THE EVALUATION FUNCTION AT UNICEF

As of December 2015, 106 country offices (82.8 percent of 128 COs) had professional posts that included evaluation in the job title, as compared to 85 in 2008, spread across all regional offices and four headquarters locations (New York, Geneva, Copenhagen and Florence). Of the 279 evaluation professionals in UNICEF, 66 percent of them were in levels-3 or above<sup>11</sup> (the number of evaluations professionals in level 3 has more than doubled since 2008). Furthermore, 79 percent (84 of 106) offices had level-3 evaluation professional in 2015, as compared to 53 in 2008. More than 94 percent of all evaluation professionals are based in field offices, indicating the highly decentralized level of the evaluation function in UNICEF.<sup>12</sup>

The Evaluation Office continues to restructure to meet its need and accountabilities towards the Strategic Plan. At the beginning of 2016, the Evaluation Office had 19 staff which included evaluation specialists in the field of emergencies, health/HIV/AIDS/nutrition, education, early childhood education, WASH, institutional effectiveness, and evaluation capacity building and knowledge management. Most of the evaluation professionals in the Evaluation Office are

of level- 4 or 5. Included in the total staff are two management positions – a Senior Adviser (P5) and a Director (D2). The Director of the Evaluation Office reports to the Executive Director, with day-to-day supervision by the Deputy Executive Director (Management) and direct access to the Executive Director. The Global Evaluation Committee advises the Executive Director on evaluation matters.

The Evaluation Office provides global leadership of the evaluation function with responsibilities in the following areas: a) governance and accountability; b) conducting evaluations; c) partnership for evaluation; d) knowledge management for evaluation; and, e) development and professionalization of the UNICEF evaluation function.<sup>13</sup>

UNICEF's evaluation activities are largely undertaken at the Regional (RO) and Country Office (CO) levels. About 90 percent of evaluations completed at UNICEF are undertaken at the CO level. The ROs, under the leadership of the regional directors, provide regional leadership of the evaluation function with reference to governance and accountability, guidance and quality assurance, conducting evaluations, partnerships for evaluation, development and professionalization of the evaluation function, and national evaluation capacity development.

### The Evaluation Policy

The revised Evaluation Policy<sup>14</sup> (which builds on the Evaluation Policy of 2008) governs the organization's evaluation function and provides a comprehensive framework for all evaluation activities undertaken by UNICEF. The policy not only lays out UNICEF's commitment to demonstrate results, transparency and accountability through an independent and

<sup>11</sup> Level-3 is the desired minimum level to ensure broad evaluation competencies, but is not required that posts be at or above that level

<sup>12</sup> Draft Annual Report on Evaluation Function -2015.

<sup>13</sup> Revised Evaluation Policy of UNICEF -2013.

<sup>14</sup> Revised Evaluation Policy of UNICEF -2013 ([http://www.unicef.org/evaluation/files/2013-14-Revised\\_evaluation\\_policy-ODS-English.pdf](http://www.unicef.org/evaluation/files/2013-14-Revised_evaluation_policy-ODS-English.pdf)).

credible evaluation system but also supports the mission, mandate and strategic priorities of the organization. Furthermore, it is aligned with the principles of the UN and with the norms and standards defined by the United Nations Evaluation Group (UNEG).

The policy provides guidance on accountabilities at the global, regional and country level in order to ensure the success of the evaluation function. The policy also prescribes performance standards for the evaluation function. Importantly, it sets out program coverage requirements for evaluation.

In UNICEF, evaluations are undertaken at the global and decentralized levels to provide evidence to inform decision-making within UNICEF and among partners. The policy recommends that evaluations should be undertaken: a) before program replication or scaling-up; b) when responding to major humanitarian emergencies; c) following long periods (5 years or more) of unevaluated program implementation; and, d) for each program outcome result component when expenditure has reached USD 10 million.

The Evaluation Policy (2013) commits the organization to an allocation of one percent of UNICEF's overall program expenditure for evaluation. In 2015, UNICEF spent 0.6 percent of its programming funds for evaluation. Although the spend on evaluation was below target, it was a significant increase compared to 0.33 percent in 2011 and 0.5 percent in 2014.<sup>15</sup>

## Quality Assurance of Evaluations

Independently assessed information on the quality of the evaluation reports produced by UNICEF is available in UNICEF's Global Evaluation Reports Oversight System (GEROS). It is an organization-wide system providing quantitative quality assessments of final evaluation reports and is managed by UNICEF's Evaluation Office. The system is augmented by quality-assurance mechanisms designed and implemented by Regional Offices and Country Offices.

The 2013 GEROS report covered 79 evaluation reports published in 2012. The report indicated that 62 percent of the evaluation reports were of good quality – rated highly satisfactory or above, an increase of 20% over 2011 reports (42 percent).<sup>16</sup> The 2014 GEROS report covered 96 evaluation reports submitted in 2013 to UNICEF's Global Evaluation Data Base. In an improvement from the previous year, the report rated 69 percent of the evaluation reports as highly satisfactory or above at an overall level.<sup>17</sup>

Finally, the 2015 GEROS report analysed 69 reports using the GEROS assessment tool which contains 58 questions/criteria.<sup>18</sup> The report noted a continuing upward trend in the observed quality of UNICEF's evaluations as reports rated as highly satisfactory or above reached the 74 percent level.<sup>19</sup>

In general, those reports that received an overall rating of less than satisfactory tended to show shortcomings in areas of methodology, recommendations, lessons learned and annexing complementary information. The limitations

---

<sup>15</sup> Draft Annual Report on the Evaluation Function 2015.

<sup>16</sup> GEROS – Global Meta-Evaluation Report 2012 – Final Report August 2013, p. i-v ([http://www.unicef.org/evaldatabase/index\\_70462.html](http://www.unicef.org/evaldatabase/index_70462.html)).

<sup>17</sup> GEROS – Global Meta-Evaluation Report 2013 – Final Report August 2014, p. i– iv ([http://www.unicef.org/evaldatabase/index\\_85763.html](http://www.unicef.org/evaldatabase/index_85763.html)).

<sup>18</sup> The GEROS assessment tool contains 58 questions distributed across six main sections (object of the evaluation; evaluation purpose, objectives, and scope; evaluation methodology, gender and human rights; findings and conclusions; recommendations and lessons learned; and, structure, logic and clarity of the report). These overarching sections are structured around fundamental areas of focus.

<sup>19</sup> GEROS – Global Meta-Evaluation Report 2014 – Final Report , September 2015, p i-iv.

in methodology and recommendations were often linked to insufficient details or omission of key components.

However, of the 90 evaluation reports (of 2015) assessed in 2016, only 53 percent (48/90) were rated highly satisfactory or above. This was a significant change compared to the upward trend of the previous three years.<sup>20</sup>

The recently concluded external assessment of the Geros found the system to be effective, but it also notes that a range of improvements are required to realize the full potential of the system to support high quality evaluation processes.<sup>21</sup> In 2013, a review commissioned by the Evaluation Office also found that the quality and coverage of evaluations needed to improve and that equity for children and gender equality were given insufficient attention in evaluations.<sup>22</sup>

### Global Level Involvement

The United Nations Evaluation Group (UNEG) is a professional network composed of the evaluation of 46 organizations across the UN system. UNEG plays an increasingly prominent role in promoting system-wide coherence around evaluation and UNICEF continues to invest in or lead UNEG activities. In 2015, UNICEF took up a leadership role in UNEG, when the Director of the Evaluation Office was appointed its Vice-Chair.

UNICEF continued to support the development and update of a global evaluation web platform known as My M&E (monitoring and evaluation), which is associated with EvalPartners. UNICEF has also continued to support the growing number of global, regional and national evaluation conferences and conclaves that help to promote evaluation capacity and coherence.

In the field of the evaluation of humanitarian action, UNICEF has played a key role in the Inter-Agency Humanitarian Evaluation (IAHE) Group and in the launch of the UNEG Humanitarian Evaluation Interest Group. In addition, UNICEF continued to work with the Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP).

## 1.7 RESULTS MONITORING AND REPORTING

Since 2011, UNICEF has invested considerable effort in modifying and strengthening its results monitoring and reporting systems. Furthermore, it has modified its organizational structure to focus on results monitoring and reporting.

### Field Results Group

The Field Results Group (FRG) was created towards the end of 2014, under a newly appointed Deputy Executive Director.<sup>23</sup> Although the focus of FRG is at regional and country level, it also works with global programs. The team of 15 staff is based at UNICEF headquarters in New York. Its mandate is to improve UNICEF's capacity in planning, implementation and demonstrating better results. The FRG has three pillars:

- **Program guidance:** which includes skill development, specifically to design result indicators and on RBM organizationally to provide clarity on expected results;
- **Performance management and accountability:** which includes reporting on performance (qualitative and quantitative), development of clear indicators to measure performance at CO level, and promoting capacity

<sup>20</sup> Information from spreadsheet compiled by Evaluation Office. It was also noted that there was change of the external firm to assess evaluation reports in 2015.

<sup>21</sup> External Evaluation of UNICEF's Geros – Report, January 2016.

<sup>22</sup> Synthesis Report on the UNICEF Evaluation Policy and Function – February 2013.

<sup>23</sup> Traditionally UNICEF had only 3 Deputy Executive Directors. This was a new position - the 4th one.

to aggregate at regional and global level, as well as reporting on the Performance Scorecard (launched in February 2016); and,

- **Implementation modalities:** to improve efficiency and effectiveness in working with implementing partners and contracting.

The Performance Scorecard incorporates qualitative information from the feedback forms received from country and regional level. This complements quantitative data obtained from the Results Assessment Module (RAM). While the primary user of this information is FRG, the data is accessible to all in UNICEF. The information on the Scorecard can be retrieved by country or by any other query. The aggregated data is also used by the Data, Research and Policy (DRP) Division for the Corporate Scorecard (data companion), which is currently being fine-tuned.

In order to strengthen the articulation of results, the FRG has introduced Strategy Notes which are developed at country level. FRG Staff report that over 85 of the Strategy Notes have been developed and each includes a theory of change, a results framework and alignment of results and strategies with expected funding. The initiative is being reviewed by Evaluation Office, Program Office, FRG and Office of Emergency Programmes (EMOPS). The FRG is also examining the feasibility and utility of incorporating result-based budgeting at country program level to strengthen relationship between results and budgets.

### Monitoring Results for Equity System (MoRES)

MoRES is in many ways UNICEF's primary operational response to monitor and measure its strategic commitment to the refocus on equity. The implementation of MoRES began in late 2011, piloting in 27 countries, before mainstreaming in late 2012. MoRES builds on the

notion that equity-focused decisions cannot be made without robust data management systems and commitment to act on the information that they generate. The MoRES framework presents four levels with a feedback loop.

To systematically identify, track and address bottle-necks that are hindering results for the most disadvantaged children, elements of the MoRES were applied in more than 80 countries in 2013, an increase from 30 countries in 2012.<sup>24</sup>

The formative evaluation of MoRES published in 2014, among other findings, noted that the coherence of MoRES with internal UNICEF systems and processes has not been fully achieved. It also noted that, while conceptually MoRES is relevant to the refocus on equity, in reality its relevance is also determined by the political, institutional and development context in the country in which it operates.<sup>25</sup>

### Results Assessment Module (RAM)

Introduced in 2013 the Results Assessment Module (RAM) presents results by geographic area. RAM provides two sets of indicators – Global Strategic Plan indicators and country specific indicators (16,000 indicators). In 2016, UNICEF has also introduced standard indicators (500) in addition to country specific indicators to facilitate aggregation.

The reporting of results is done through the RAM reporting portal. Each manager reports two times a year which includes a 500-word description and traffic lights on progress. The RAM annual reporting also includes strategic monitoring questions based on the outcome area selected.

### Humanitarian Performance Monitoring

Within the wider results monitoring and reporting framework at UNICEF, the Office of Emergency Programmes (EMOPS) provides

---

<sup>24</sup> UNICEF Annual Report 2013.

<sup>25</sup> Formative Evaluation of UNICEF's MoRES – August 2014.

ongoing support to country and regional offices for Humanitarian Performance Monitoring (HPM) as first fully rolled out in 2012. In 2015, for example, EMOPS HPM specialists were deployed to country and regional offices to provide direct support for a total of 180 days. These included missions for the Ebola crisis, Nepal, Yemen and the European migration crisis. Country offices also received remote support on 39 activities related to setting up and managing HPM systems. Work continued on the refinement of HPM guidance and tools drawing from country experience.

Also in 2015, the development of information management platforms for HPM progressed and was mainstreamed into the wider corporate eTools project spanning a range of web-based platforms and applications to increase work process efficiencies in development and humanitarian contexts. Through the end of 2015, nine country offices with different levels and types of humanitarian action were engaged in testing and using an information management base for future HPM-related products planned for release in 2016.

### Annual Report of Executive Director

The Executive Director of UNICEF reports annually on the progress made and results achieved against the MTSP focus areas (up to the period ending in 2013) and on the Strategic Plan outcomes (from 2014). This report also summarizes UNICEF's progress related to various mandates of General Assembly resolution 67/226 on the QCPR of operational activities for development of the UN system. This report also combines information on global trends and country specific indicators relevant to UNICEF's mandate.

The annual report also addresses developments in the implementation of program principles and strategies. The 2015 report (on 2014 programming) highlights: UNICEF's capacity development support to communities and governments at national and local levels;

equity-focused and evidence-based advocacy; expansion of partnerships; South-South cooperation; and, innovations in programming.

The annual report is supported by a large data companion report which tracks a significant number of key performance indicators across the seven outcome areas and the two cross-cutting themes of humanitarian action and gender equality.

### Evaluability of the Strategic Plan

The recently completed evaluability assessment found that the Strategic Plan 2014-2017 was widely viewed as the most coherent and technically sound plan that UNICEF has developed to date. However, the assessment noted that a significant challenge lies in the lack of explicitly stated logical connections of the causal chain from outputs to outcomes and from outcomes to impact. While results statements are found at each level of the Theory of Change, there is very little explanation of how one level leads to or contributes to another. Furthermore, it also notes that, while indicator definitions were generally sound, there is scope to improve some performance indicators. The assessment also noted that, while the entire Strategic Plan is developed with an equity lens, the approach to measurement of equity is weak.



## CHAPTER 2

# METHODOLOGY

## CHAPTER 2

This Chapter presents a brief overview of the methodology used in the review. A more complete description of the review methodology is provided in Annex C.

### 2.1 EVALUATION POPULATION AND SAMPLE

The aim of the sampling process was to identify a sample of approximately 70 evaluation reports which would be illustrative of UNICEF's effectiveness in humanitarian and developmental programming during the period 2012 to 2015. With a review encompassing approximately 70 evaluation reports, experience indicates that almost all of the effectiveness criteria and sub-criteria will be covered by a sufficient number of evaluations to allow for a credible statement on the distribution of findings across the four levels used by the review.

The review identified a sample frame of 263 evaluation reports uploaded to the UNICEF evaluation data base in the period under review. This was done after excluding reports with a GEROS rating of unsatisfactory and reports which were unlikely to address the key evaluation criteria used in the study. The review then developed a modified random sample of 73 evaluation reports for submission to an internal evaluation quality review process (Annex E). As a result of that process, seven reports (9.6 percent) were excluded from the analysis due to quality scores which were lower than the minimum score determined in advance (25 of 40 available points).

Interestingly, the rate at which evaluation reports were excluded for quality reasons was higher for the 2015 DER than for the one undertaken in 2012. During the previous DER, four of 66 evaluation reports (6.0 percent) were excluded for failing the quality review. The quality review criteria, the individual weighting scores for each criterion and the detailed

results of the quality review process, including the distribution of total scores, are all presented in Annex E.

At the same time, it is important to note that the team of analysts (including the team leader and deputy team leader) which reviewed the 66 remaining evaluations in the sample noted that those evaluations which did qualify for review were generally more clearly written, more consistently well structured and, as a group, were much more specific in their coverage of the review's evaluation criteria.

The final sample reviewed for this DER included 66 evaluation reports (25 percent of those in the sample frame of 263). It has the following characteristics illustrated in detail in Annex D:

- The distribution of evaluation reports in the sample is very similar to the distribution of reports by **year of publication** in the overall population;
- The distribution of sampled reports by UNICEF Program region is very close to the regional distribution of reports in the sample frame;
- The sample tracks very closely to the distribution of UNICEF evaluation reports by level of country development;
- In order to better represent the budgetary balance of funding across defined output areas of UNICEF programming, the sample slightly over-represents the health sector with 21.2 percent of the sample versus 12.5 percent of the population.

In summary, the resulting sample of 66 evaluation reports (Annex C) represents a **modified (and thus purposive), stratified random sample**

of approximately 25 percent of the evaluation reports included in the sample frame. It is illustrative of the results of UNICEF's efforts to assess development effectiveness during the period 2012 to 2015. The sample presents a faithful reflection of UNICEF's own findings and conclusions regarding the effectiveness of its development programs as evaluated in the field.

However, as noted in the limitations section (2.4), the sample does not provide sufficient coverage of evaluations of humanitarian programming to allow the review to make conclusions or observations on UNICEF's effectiveness in humanitarian programming.

Not all of the evaluation reports reviewed contained information on program expenditures during the evaluation period. In fact, 46 of the 66 evaluation reports in the sample included information on UNICEF expenditures under the program being evaluated. For these evaluations, UNICEF expenditures amounted to USD 1.129 billion.

## 2.2 CRITERIA

As with all DERs, the review methodology focused on six main criteria and 20 sub-criteria as the core elements of effective developmental and humanitarian programming.

**TABLE 1** Criteria for Assessing Developmental and Humanitarian Effectiveness

<b>1. Relevance</b>
1.1 Programs are suited to the needs of target group members
1.2 Programs are aligned with national humanitarian and development goals
1.3 Effective partnerships with government, civil society and humanitarian and development partners
<b>2. Achievement of Development and Humanitarian Objectives and Expected Results</b>
2.1 Programs and projects achieve stated objectives
2.2 Positive benefits for target group members
2.3 Substantial numbers of beneficiaries/contribution to national humanitarian and development goals
2.4 Significant changes in national development policies/programs
<b>3. Cross Cutting Themes: Inclusive Development and Humanitarian Which can be Sustained (Gender Equality, Environmental Sustainability and Equity)</b>
3.1 Programs effectively address gender equality
3.2 Changes are environmentally sustainable
3.3 Programs effectively address the issue of equity
<b>4. Sustainability</b>
4.1 Program benefits are likely to continue
4.2 Programs support institutional and community capacity
4.3 Programs strengthen enabling environment for humanitarian and development
<b>5. Efficiency</b>
5.1 Program activities are cost efficient
5.2 Programs are implemented/objectives achieved on time
5.3 Systems for program implementation are efficient
<b>6. Using Evaluation and Monitoring to Improve Humanitarian and Development Effectiveness</b>
6.1 Systems and processes for evaluation are effective
6.2 Systems and processes for monitoring are effective
6.3 Results-based management systems are effective
6.4 Evaluation results used to improve humanitarian and development effectiveness

One new sub-criterion, not used in the 2012 review, has been added to improve the relevance of the review to UNICEF's current programming. Sub-criteria 3.3 examines the extent UNICEF supported programs effectively address issues of equity and, thereby, respond to a key strategic direction for the organization.

This new sub-criterion serves to extend the analysis and enrich the information gathered from the reviewed evaluation reports. It does not diminish the comparability of results across the two completed DERs, because all 19 sub-criteria from the 2012 study were used in the current round and one additional criterion was added.

## 2.3 REVIEW PROCESS AND DATA ANALYSIS

Each evaluation report in the sample was reviewed by a member of a small review team of two reviewers and two senior members (including the team leader). The reviewers classified the evaluation findings in each evaluation report as they apply to the six evaluation criteria and 20 sub-criteria presented in Table 1. Each finding was classified using a web-based evaluation review reporting tool under one of five different headings:

1. Not Addressed
2. Highly Unsatisfactory
3. Unsatisfactory
4. Satisfactory and
5. Highly Satisfactory

The draft final report combines the quantitative evidence from the 66 evaluation reports reviewed, with qualitative evidence gathered from the review of a sample of thematic evaluations and from interviews with UNICEF staff. This approach is entirely consistent with the data gathering and analysis process used in the 2012 review and in all other DERs completed since 2010.

## 2.4 LIMITATIONS

As a meta-synthesis of previously completed evaluation reports, this development effectiveness review has been subject to a set of limitations. The most important of these include:

- **Sampling Bias:** The evaluation sample, while it began with a stratified random sample, was adjusted to purposely improve representation of evaluations in the health sector and evaluations from countries in sub-Saharan Africa. The resulting purposive sample is not a statistically valid random sample of UNICEF evaluations which can be used to calculate confidence intervals;
- **Bias in the Evaluation Universe:** Any biases in the universe of UNICEF evaluations (when compared to programs) can be expected to be reflected, to some degree in the evaluation sample, despite efforts to balance representation, in the evaluation sample;
- **Limits to Stratification:** The evaluation sample has been balanced to be largely reflective of the distribution of evaluation reports produced by UNICEF in relation to geographic area, year of publication and outcome area of the strategic plan. However, the resulting strata of the sample are too small to allow for meaningful reporting of results in each sub-component (the different outcome areas for example). Nonetheless, it has been possible to compare some results across broader aggregations of stratification (for example, between least developed countries (LDCs) and other categories of countries);
- **Coverage of Humanitarian Programming:** Limits to stratified analysis in the sample have particular importance when considering the humanitarian dimension of UNICEF's work. Because evaluations of humanitarian programming accounted for only 2.7 percent of all UNICEF published evaluations during the review period, a very small number could be included in the review if the sample

was to reflect the overall population. While humanitarian focused evaluations were somewhat oversampled (they made up 6.1 percent of the sample), they still were too limited in number to allow any generalization of results to UNICEF's overall experience in humanitarian programming. It is important to note that the evaluation universe, and the sample, both include evaluations classified under other output areas, for example water and sanitation or health programming, which take place in a humanitarian or emergency contexts and so have potential implications for humanitarian programming. However, the actual review process of evaluations included in the sample, did not provide enough evidence of finding on humanitarian operations to support conclusions on humanitarian programming;

- **The Retrospective Nature of Meta-Synthesis:** Evaluation reports are, by definition, retrospective in that they provide ex-ante assessments of the effectiveness of programs. Many of the programs evaluated from 2012 to 2015 originated some years before they were evaluated. Further, publication of the findings may only have occurred some time after the evaluation field work was completed. This means that the review must take care not to hold older programs accountable for new strategies or to generalize historical results to current practice;

While these limitations are real, none strongly undermine the utility of the results of the development effectiveness review process. The process provides UNICEF with one of the very few approaches available for aggregating field level, ground proven findings on effectiveness. Further, it allows UNICEF to take advantage of investments in evaluation which have been already made, by organizing and analyzing the results of those investments in a systematic way.



## CHAPTER 3

# FINDINGS OF UNICEF'S DEVELOPMENT EFFECTIVENESS

## CHAPTER 3

This Chapter presents the results of the review as they relate to five of the six main criteria and associated sub-criteria used to assess effectiveness.

Because of the importance of evaluation and monitoring issues, those are addressed in Chapter Four. For each sub-criterion the report presents first the extent it was addressed in the evaluation reports reviewed (coverage). It then presents the findings of the DER with respect to each sub-criterion – including both the quantitative results of the review of evaluations and the qualitative results of interviews and the review of thematic evaluations carried out separately. The chapter also includes a discussion of the factors which contributed to, or detracted from, the achievement of a positive finding for the sub-criteria.

Where numbers warrant a comparison among different types of evaluations, these comparisons are made. For the most part, this is limited to a comparison of results by the level of development of the countries where UNICEF supports activities and programs: Lower Income (LIC), Lower Middle Income (LMIC), and Upper Middle Income (UMIC) countries. This was the only grouping in the 2016 review with sufficient numbers of evaluations in each group to warrant any meaningful comparison of results. Of the 66 evaluations in the sample, 23 evaluations addressed programming in LICs, 26 in LMICs and 17 in UMICs. With seven UNICEF output areas and a sample of 66 evaluations (in contrast to three thematic areas in the

2012 review) it was not possible to make valid comparisons of results by output area in the 2016 Review (Annex C).

Finally, in each section the report highlights any significant differences reported in this Development Effectiveness Review when compared with the findings from the 2012 Review (Annex I). Since many sub-criteria report very similar levels of positive and negative findings across the two different reviews, the report only points out the most significant differences. If no mention of difference is made when discussing a specific criterion, the reader can conclude that the proportion of positive findings (satisfactory or highly satisfactory) is no more than five percent different across the 2012 and 2016 reviews.

In reporting on the frequency with which contributing factors are cited the report (as in 2012) makes use of the terms “most”, “many”, “some” and a “few” to describe the number of evaluations referring to a given factor (See Box 3). In addition, the order in which contributing factors are presented reflects the relative frequency of their citation in the reviewed evaluation reports.

Table 2 summarizes coverage of each sub-criterion (with the letter n column reporting the number of evaluations that addressed each sub-criterion).

### BOX 3

#### Frequency of Factor Observations

**Most** = over three quarters of the evaluation covering the criteria

**Many** = between one half and three quarters

**Some** = between one fifth and one half

**Few** = less than one fifth

**TABLE 2** Coverage by Sub-Criteria: 2016 and 2013

Criteria and Sub-Criteria	n* 2016	Coverage Level 2016 †	n 2013	Coverage Level 2013
<b>Relevance</b>				
1.1 Programs suited to the needs and/or priorities of the target group.	56	Strong	60	Strong
1.2 Programs align with national development goals.	56	Strong	61	Strong
1.3 Effective partnerships with governments, bilateral and multilateral development and humanitarian organizations and NGOs.	46	Strong	56	Strong
<b>Objectives Achievement</b>				
2.1 Programs and projects achieve their stated humanitarian and development objectives.	63	Strong	61	Strong
2.2 Programs and projects have resulted in positive benefits for target group members.	64	Strong	62	Strong
2.3 Programs and projects made differences for a substantial number of beneficiaries.	58	Strong	61	Strong
2.4 Programs contributed to significant changes in national humanitarian and development policies and programs.	55	Strong	43	Moderate
<b>Cross Cutting Themes</b>				
3.1 Extent UNICEF support addresses gender equality.	58	Strong	42	Moderate
3.2 Extent to which changes are environmentally sustainable.	12	Weak	15	Weak
3.3 Programs effectively address issues of equity.	54	Strong	N/A	N/A
<b>Sustainability</b>				
4.1 Benefits likely to continue after project or program completion.	56	Strong	59	Strong
4.2 Projects and programs are reported as sustainable for institutional and/or community capacity.	56	Strong	59	Strong
4.3 Programming contributes to strengthening the enabling environment for development.	48	Strong	40	Moderate
<b>Efficiency</b>				
5.1 Program activities are evaluated as cost/resource efficient.	52	Strong	48	Strong
5.2 Implementation and objectives achieved on time.	39	Moderate	39	Moderate
5.3 Systems and procedures for project/program implementation and follow up are efficient.	42	Moderate	35	Moderate
<b>6. Evaluation</b>				
6.1 Systems and process for evaluation are effective.	56	Strong	42	Moderate
6.2 Systems and processes for monitoring and reporting are effective.	60	Strong	59	Strong
6.3 Results-based management systems are effective.	61	Strong	24	Weak
6.4 Evaluation is used to improve humanitarian and development effectiveness.	66	Strong	61	Strong

\* n = number of evaluations addressing the given sub-criterion

† Strong: n = 46 – 66 • Moderate: n = 31 – 45 • Weak: n = less than 31

In general terms, the level of coverage across the criteria and sub-criteria of the review improved in the 2012 to 2015 period, when compared to the earlier review which examined evaluations from 2009 to 2012.

In the 2012 review, 12 of 19 sub-criteria were rated strong, five were rated moderate and two were weak in coverage. The criteria with weak coverage were 3.2: environmental sustainability and 6.2: results-based management systems are effective. By the 2016 review, 18 of 20 sub-criteria (with one new additions) were rated strong and only one, 3.2: environmental sustainability was rated weak. Particularly notable is the fact that coverage of sub criterion 3.1 on how programs and projects address gender equality has moved from only moderate coverage in the 2013 DER, to strong coverage (58 of 66 evaluation reports) in 2016.

### 3.1 RELEVANCE

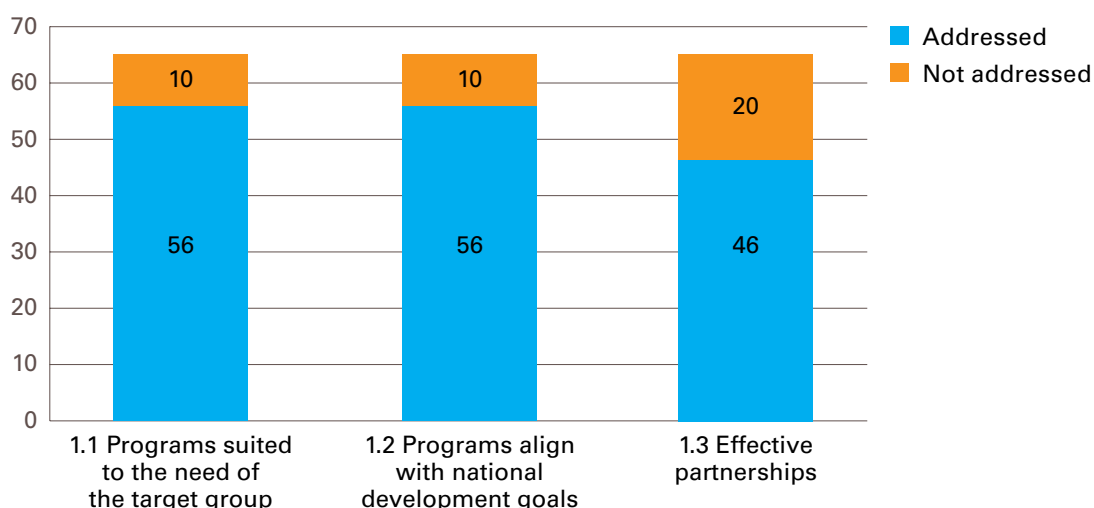
#### 3.1.1 Coverage

The review breaks down the criteria of relevance into three specific but inter-related positive statements on UNICEF programming:

- 1.1 UNICEF supported programs are suited to the needs and/or priorities of the target group.
- 1.2 UNICEF programs align with national development goals.
- 1.3 UNICEF has developed effective partnerships with governments, bilateral and multilateral development and humanitarian organizations and NGOs.

As in the 2012 review, coverage of all three sub-criteria relating to relevance was rated as strong. In particular, the evaluation reports reviewed very often addressed sub-criterion 1.1 “Programs are suited to the needs of the target group” and 1.2 “Programs align with national development goals”. A smaller number of evaluations (46) addressed sub-criterion 1.3 on UNICEF’s engagement in “Effective Partnerships”. This pattern of coverage is similar to the pattern observed in the 2012 DER for UNICEF, which also saw fewer evaluations address sub-criterion 1.1 on effective partnerships, but with all three rated strong in coverage in both reviews.

**FIGURE 3** Number of Evaluations Addressing Sub-Criteria for Relevance



### 3.1.2 Key Findings: Relevance

As a group, the evaluations reviewed for the current DER found that programs supported by UNICEF were rated as relevant across all three sub-criteria.

A full seventy percent of the evaluations reviewed reported findings that UNICEF programs were suitable to the needs of the target group (sub-criterion 1.1) at either the satisfactory or highly satisfactory level. There were no evaluations that rated sub-criterion 1.1 as highly unsatisfactory. While this is a strongly positive result, it is notably lower than the ninety percent of evaluation reports with findings on sub-criterion 1.1 which were rated satisfactory or highly satisfactory in the 2012 review.

The reported results were most strongly positive with regard to sub criterion 1.2 with 96 percent of the evaluations reporting satisfactory or highly satisfactory findings regarding the alignment of UNICEF supported programs and activities and national development goals. This finding is consistent with the results of the 2014 global evaluation of UNICEF's Upstream Work in Basic Education and Gender Equality. That found that upstream work was highly relevant in terms of its consistency with beneficiaries' requirements, country needs, global priorities and partner and UNICEF policies.<sup>26</sup>

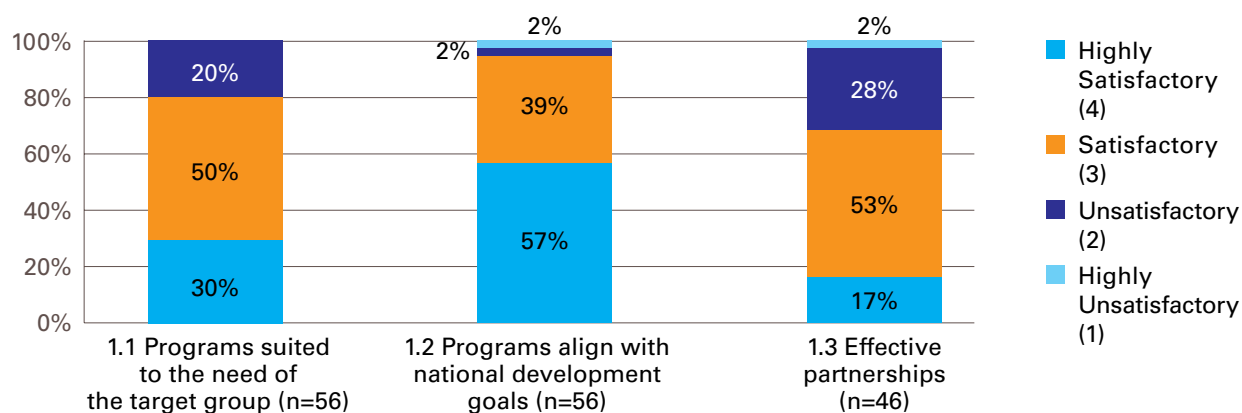
#### BOX 4 Meeting Beneficiary Needs Through Consultation on Messaging

"the communities are very much satisfied with hygiene promotion activities and confirmed that the hygiene messages are very relevant to their situation. This is mostly due to the fact that the program's approach in terms of planning and implementation has been built on consultation with sector stakeholders and partners and takes into consideration previous studies together with the culture and traditions of the targeted community, which ensures that messages are culturally appropriate and relevant."

*Program Evaluation of UNICEF Assisted Water, Sanitation and Hygiene Program in Sudan. 2012, p.42.*

While still a positive result, findings on sub-criterion 1.3, the extent that UNICEF has engaged in successful partnerships with government, development partners and civil society were in the same range as sub-criterion 1.1, with 70 percent of reviewed evaluations reporting satisfactory or highly satisfactory findings.

**FIGURE 4 Findings for Relevance**



<sup>26</sup> UNICEF's Upstream Work in Basic Education and Gender Equality – 2003 to 2012. (20014). P.8.

As with sub-criterion 1.1, the findings on effective partnership (sub-criterion 1.3) rated positively somewhat less often in the 2016 review (seventy percent), when compared with the 2012 review (87 percent).

It is worth noting that the sample of evaluations reviewed in 2012 was more strongly focused on Least Developed Countries (LDC) than the 2016 sample. In 2012, over half (55 percent) of the evaluations focused on programs in LDCs, while in 2016, the proportion of Low Income Countries (Annex C) was just 34.8 percent. Since both samples were balanced in relation to the population of evaluations, this would seem to confirm that UNICEF support and the evaluations it carries out are now more balanced across the spectrum of country classifications in which it works.

Since sub-criterion 1.3 on effective partnerships concerns UNICEF's relationships with host governments, multilateral development partners and civil society organizations, a shift to policy engagement with national government partners (with perhaps less emphasis on partnerships with Non-Government Organizations (NGOs) as implementing partners) might also, to some degree, explain the relatively lower proportion of positive findings among the evaluations reviewed in 2016.

Results for sub-criterion 1.3 on effective partnerships were more often positive when programs being evaluated occurred in Upper-Middle

Income Countries (UMIC) as illustrated in Annex G. For programs in UMICs, 75 percent of reviewed evaluations found that UNICEF was engaged in effective partnerships at the satisfactory or highly satisfactory level.

The global evaluation of UNICEF's work in protecting children from violence also pointed to the agency's effectiveness in mobilizing partnerships in order to contribute to interventions and initiatives.<sup>27</sup>

Specific examples of relevance from the evaluation reports include:

- Building program implementation strategies and approaches based on consultations with stakeholders, partners and communities resulted in a stronger program more suited to target group needs (Sudan and many others);
- Communicating the program design nationally and locally contributed to recommended changes in design which encouraged greater ownerships (El Salvador);
- Combining geographic targeting of poorest-served areas and demographic targeting of underserved populations helps promote equity and limit the risk of exclusion (Niger);
- Government involvement in the development of the program plan improved alignment with national priorities (India and Lesotho);

## BOX 5

### Contributing to Partnerships in Maternal Health in Argentina

"UNICEF's contribution took the form of actions in various areas, from financial contributions to stakeholder coordination and capacity building. Two years after its implementation, UNICEF was able to form a multi-sectoral, collaborative partnership, with participation of national and provincial governments, academic institutions, management personnel and members of the health teams of 95 public maternity hospitals in 9 priority provinces."

*Evaluación de la Iniciativa Maternidades Seguras y Centradas en la Familia 2010-2014. 2015. P.30*

<sup>27</sup> Protecting Children from Violence – A Comprehensive Evaluation of UNICEF's Strategies and Program Performance (2015).

- Strong collaborative relationships between the lead government agency (the Ministry of Health) and multilateral agencies, as well as bilateral development partners can play a key role in the success of UNICEF programming (Malawi).

### 3.1.3 Contributing Factors

#### Positive Contribution Factors: Relevance

Of course, a wide range of factors made a positive contribution to the relevance of UNICEF-supported programs.<sup>28</sup> Most evaluations made specific reference to the positive alignment between UNICEF programming and national development plans and programs, including: overall national development plans, national anti-poverty action plans and national vision statements.

Many evaluations also pointed to the strong link between UNICEF supported programs and national plans, programs and strategies at a sectoral level. These included national sector strategies in health, WASH, education, juvenile law, gender-based violence, and HIV/AIDS.

Important factors promoting relevance noted in some evaluations included:

- Engaging key stakeholders, including representatives of target group members in meaningful consultations on program design, so that the program outputs reflect and meet the needs of the target group and are supported by stakeholders;
- The use of a systematic process of needs assessment including surveys, epidemiological studies, poverty and disease burden mapping, and formal situation analysis exercises;
- The use of specific targeting measures to address the needs of groups which have been excluded on either geographic or

social terms, combined with the use of data on access to services and needs by different excluded segments of the population; and,

- Effective mechanisms for coordinating UNICEF support in a given sector of action with the actions of governments, multilateral agencies, bilateral donors and NGOs.

#### A few evaluations noted other positive factors contributing to relevance:

- UNICEF's ability to engage with different networks established to provide technical inputs to sectoral policies and programs;
- UNICEF's technical capacities and its reputation for technical strength in specific areas such as mother and child health services and child protection; and,
- UNICEF's ability to engage on matters relating to international conventions including the UN Convention on the Rights of the Child, the Convention to Eliminate All Forms of Discrimination Against Women, and the Convention on the Rights of Persons with Disabilities.

#### Negative Contributing Factors: Relevance

The group of evaluation reports with negative findings on relevance was relatively small but, of that group, a significant number identified factors which reduced the relevance of programs and thus made a negative contribution.

Many pointed to the inadequacy or the apparent absence of a systematic assessment of the needs of the target group members as an important factor undermining relevance. Many evaluations also identified weak engagement and lack of participation by host governments or other key stakeholders as factors undermining program relevance.

<sup>28</sup> The different positive and negative contributing factors are presented in the report in the relative order of the frequency with which they were identified in the evaluation reports. As noted in Box 3, they are referred to as: "most" found in over three-quarters of the relevant evaluations; "many" between half and three quarters; "some" between one fifth and one-half; and, a "few"; less than one fifth.

Some evaluation reports identified as negative factors:

- Poor targeting of interventions and their outputs so that benefits bypassed the poor and other identifiable vulnerable populations; and,
- Weak and generally ineffective coordination mechanisms, often in combination with unclear roles among partners.

A few evaluations note that poor partner capacity, with a limited ability to engage in collaborative program planning, also limited the relevance of the agreed program of UNICEF support.

### 3.1.4 Summary and Implications: Relevance

In summary, as in the 2012 study, the evaluations reviewed found UNICEF's programming to be relevant to the needs and priorities of both countries and individuals. While findings are less positive than in 2012 for sub-criteria 1.1 on meeting target group needs and 1.3 on effective partnerships, they remain strongly positive. It might seem an anomaly that even four percent of the sample evaluation reports would include an unsatisfactory finding for sub-criterion 1.2, on the alignment of UNICEF support with national development goals. It is important to remember that this represents just two of 56 evaluation reports which addressed this sub-criterion. The two evaluation reports in question either explicitly noted that the UNICEF supported program was not aligned with national plans (Paraguay 2013) or identified the need for closer alignment between UNICEF programming and the national program (Ethiopia 2014).

It is also important to note that the 70 percent of evaluation reports finding that UNICEF programs meet the needs of target group members is somewhat inconsistent with the findings for sub-criterion 2.2 on the extent that UNICEF programming provides benefits for target group members. For that sub-criterion 95 percent of evaluations reported positive

findings. So even where evaluations do report design issues relating to the fit between programs and target group needs, they almost always indicate that there are substantial benefits for the same group.

An overview of the positive and negative factors contributing to the relevance of UNICEF programs suggests there is room to improve the design of a significant number of programs by:

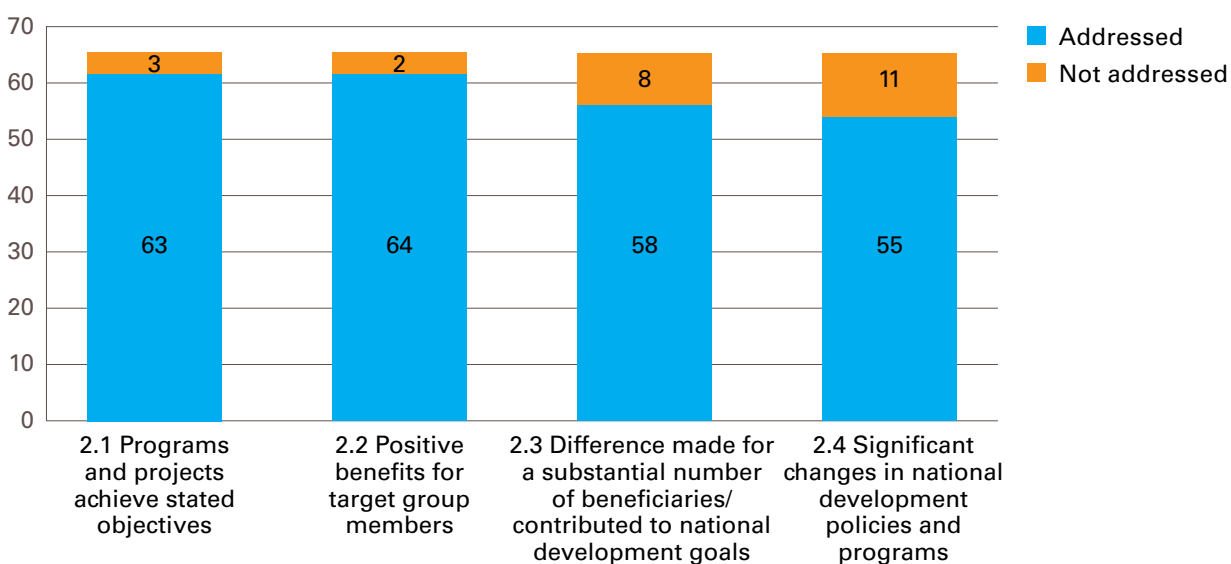
- More often engaging key stakeholders, including representatives of target group members in meaningful consultations on program design;
- Making more use of systematic processes of needs assessment;
- Using specific and explicit measures to identify and meet target group members needs while making use of data on access and use of services by the same members; and,
- Investing more effort in the development and operation of processes for coordinating UNICEF support with the actions of governments, UN agencies, bilateral agencies and NGOs.

## 3.2 ACHIEVEMENT OF OBJECTIVES

### 3.2.1 Coverage

There are four sub-criteria which make up the review's examination of how well UNICEF supported programs achieve their stated objectives:

- 
- 2.1 Programs and projects achieve their stated humanitarian and development objectives.
  - 2.2 Programs and projects have resulted in positive benefits for target group members.
  - 2.3 Programs and projects made differences for a substantial number of beneficiaries.
  - 2.4 Programs contributed to significant changes in national humanitarian and development policies and programs.
-

**FIGURE 5** Number of Evaluations Addressing Sub-Criteria for Objectives Achievement

There is strong coverage across all four sub-criterion related to how UNICEF supported programs achieve their stated objectives. This is a slightly better coverage pattern than the one achieved for the 2012 DER of UNICEF, when sub-criterion 2.4 (regarding how UNICEF actions contribute to significant changes in national development plans and priorities) received only a moderate level of coverage.

The higher level of coverage of sub-criterion 2.4 in the 2016 review could reasonably be attributed to UNICEF programming in the 2012-2015 shifting towards a **greater emphasis on “upstream” work through greater engagement in the development of national policies, programs and guidelines**. In turn, there has clearly been a higher level of attention to UNICEF’s contribution to national development policies and guidelines reflected in evaluations carried out in the past three years than in those completed from 2009 to 2011.

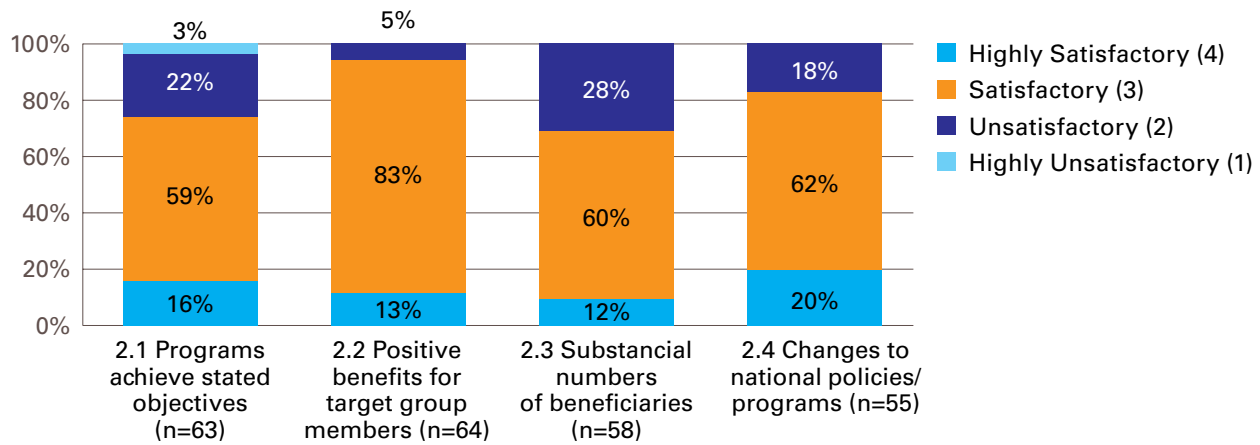
### 3.2.2 Key Findings

The overall findings for the key evaluation criterion of objectives achievement are very positive for all four sub-criteria. All sub-criteria, except for 2.3 which assesses the level of program reach in terms of the numbers of beneficiaries reached were positively rated (satisfactory or highly satisfactory) in more than 75 percent of the evaluations reviewed, and it only missed that level by three percent. **This is a strongly positive result.**<sup>29</sup>

In comparison to the results of the 2012 review, findings for objectives achievement were somewhat more positive. The proportion of evaluations with findings rated satisfactory or highly satisfactory was virtually identical (within two percentage points) for sub-criterion 2.1 on achieving the program’s stated objectives and sub-criterion 2.3 on reaching substantial numbers of beneficiaries.

<sup>29</sup> It is important to note here that the review team has made very effort to ensure that the classification of evaluation findings across the four levels from highly satisfactory to highly unsatisfactory is as objective as possible. For the criteria used to classify each evaluation finding see Annex J, the analysis guide. For a description of quality assurance processes, see Annex C on methodology.

**FIGURE 6 Findings for Objectives Achievement**



However, the proportion of evaluations reporting positive findings was notably higher in the 2016 review in relation to sub-criterion 2.2, positive benefits for target group members than in the 2012 review (95 percent in 2016 versus 86 percent in 2012). Positive findings were also somewhat more frequent in 2016 for sub-criterion 2.4 on UNICEF’s contribution to changes in national policies with 82 percent of evaluations reporting positive findings in 2012 versus 77 percent in 2012. There seems to be a modest but notable trend to continued improvements in how UNICEF supported programs are able to achieve their stated objectives.

Comparing evaluation findings for objectives achievement across different country types provides some interesting perspectives.

The proportion of evaluation reports with positive findings (satisfactory or highly satisfactory) in relation to sub-criterion 2.1 on achieving stated program objectives is much higher for programs in LICs (85 percent) and LMICs (80 percent) than in UMICs, where it drops to just 53 percent. This may well reflect the relative difficulty of setting realistic, time bound targets in the area of policy engagement and advocacy which is the predominant form of UNICEF engagement in UMICs.

Similarly, the reviewed evaluations reported positive findings relating to sub-criterion 2.3 on reaching substantial numbers of beneficiaries much less frequently for programs in UMICs (47 percent) than for those in LICs (80 percent) and LMICs (83 percent). This could be a reflection of the indirect nature of benefits that would potentially flow to large numbers of beneficiaries as a result of engagement in advocacy and national policy development. It could also be a reflection of the significant time lags normally experienced between policy engagement, the development of new national policies, and the benefits eventually experienced by target group members.

**BOX 6 Improving Child Nutrition in Ethiopia**

“Data from the first three years of the program from evaluation surveys and the Community-based Nutrition Program (CBN) routine data, showed that the program exceeded its stated aims, both in terms of implementation and outcomes. Indicators of infant and child feeding practices and nutritional status objectives were met or exceeded.”

*Assessing the impact on child nutrition of the Ethiopia Community-based Nutrition Program. 2013.p.39*

### Types of Positive Results Reported

Many evaluation reports found that UNICEF supported programs and projects had achieved their stated objectives to the outcome level while some reported only to the output level. The most frequently noted types of objectives achieved, and subsequent benefits experienced by target group members (including service providers) included:

- Improved levels of skills, positive changes in behaviour, and higher levels of confidence among service providers with subsequent gains in quality of service and community access were noted in many evaluation reports;
- Strengthening national policies, including better coordination across different sectors such as water and health was a noted positive benefit of UNICEF programming in many evaluation reports. Examples include policies on refugee education (Turkey), laws on child protection (Montenegro), laws on juvenile justice (Bosnia), policies on humanitarian assistance (the DRC), guidelines on social behaviour change to prevent violence against women (South Africa), national WASH guidelines (Sudan) and national policies on integrated health services (Ghana);
- Strengthened community structures capable of assessing needs and taking action to improve services and access reported in some evaluations;
- Improvements to physical infrastructure of schools and health facilities (Madagascar, Tajikistan, Jordan, Mali, Moldova, Namibia);
- Improved access to better water and sanitation facilities (Burundi, Sudan, Indonesia, Laos PDR, Pakistan, Nigeria, Mozambique);
- Sustaining the implementation of national policy by augmenting national policies and programs with resources or technical support (Burkina Faso, Pakistan, Madagascar, Jordan, Malawi, Mongolia, Indonesia);
- Effectively addressing poverty through a multi-sectoral approach (including cash grants) and women's economic empowerment (Colombia, Somalia, South Africa, Lesotho);
- Improved participation rates and educational outcomes, particularly for girls (Benin, Jordan, Madagascar, Palestine, Turkey);
- Improved participation and confidence for girls taking part in sports with an associated improvement in outcomes at school (Namibia);
- Improved demand for strengthened health services and better health seeking behaviour (Benin, Burkina Faso, Djibouti, Ghana, Malawi, Togo);
- Improvements in health outcomes including maternal mortality ratios, improved nutritional status and reduced incidence of malaria (El Salvador, Ethiopia, Ghana, India, Zimbabwe);
- Improved psychosocial well being of refugee children (Jordan);
- Better handling of legal cases involving children by a better qualified judiciary (Tajikistan) and,
- Improved social protection for children made vulnerable by food insecurity (Niger).

The above list is not exhaustive but it does illustrate the contribution made by UNICEF programming to a very diverse set of positive results as reported in the evaluations reviewed.

### 3.2.3 Contributing Factors

#### Positive Contributing Factors: Objectives Achievement

The evaluations analyzed for the 2016 review cited a wide range of both positive and negative contributing factors. On the positive side, many evaluations noted as positive factors:

**BOX 7****Improved Children's Well-Being in Syrian Refugee Camps in Turkey**

"Results from focus group discussions with boys and girls aged 9-12 and male and female adolescents aged 13-18 revealed that attending Child Friendly Schools (CFS) had the biggest effect on children's emotional well-being; effectiveness in this aspect scored 72 out of the 100 points available. Within this category, "mood" (feeling happier, not angry, sleeping better, getting out of isolation) scored 80%."

*UNICEF's Emergency Psychosocial Support for Syrian Children in Jordan: An Evaluation. 2015. P.3.*

**BOX 8****Reaching Girls through Soccer in Namibia**

"Overall, the program had made a huge difference in the lives of girls. Girls at all sites visited spoke spontaneously of how the program had changed their lives and had helped them to become team players and to develop a positive attitude to their school work. The program motivated them to pursue their dreams and prevented them from getting involved in risky behaviour. But perhaps most importantly, the program had helped them to develop self-confidence, self-esteem, self worth and self-efficacy. The program also helped the girls to develop leadership skills."

*Impact Evaluation of the Namibia Football Association's Galz and Goals Sports for Development Program. 2015. P.6.*

- Well-designed training initiatives aimed at improving the skills, confidence and, especially, the attitude of service providers in health, education, justice, policing and refugee services toward target group members, especially adolescents and young girls; and,
- Willingness on UNICEF's part to engage with and support national authorities in the development of national policies and programs in order to promote national ownership, along with a greater level of commitment on UNICEF's part to policy engagement and advocacy.

Some evaluations pointed to:

- The inclusiveness of program planning and review processes and a willingness to engage with key stakeholders, especially at community level during program planning; and,

- Innovations and innovative models for the design and implementation of projects and programs such as improved teaching models (Benin), use of mobile teams (Mongolia), use of radio programming (Mongolia), alternative family based service delivery (Montenegro) and, distance learning techniques (Zambia).

A few evaluations reported each of the following as positive factors contributing to objectives achievement:

- UNICEF's capacity to deploy resources quickly to intervene in emerging humanitarian situations; and,
- Appropriate measures to identify target group members and to structure service delivery to ensure participation by the target group.

### Negative Contributing Factors: Objectives Achievement

Negative factors which limited the success of UNICEF programming in achieving its stated objectives were naturally cited less frequently in the evaluations reviewed since overall, most reported positive findings for this criteria. However, a few evaluations noted as negative factors:

- General weaknesses in program/project design, especially relating to lack of clarity in design and weak linkages from activities to outputs and outcomes. These weaknesses extended to a general lack of targets and weak monitoring and evaluation systems;
- Absence of sufficient baseline and/or ongoing data on target group needs to support effective design;
- Poor or inadequate targeting of vulnerable populations, either due to lack of information on where the vulnerable group members were located and their specific needs or relatively weak measures to ensure their participation;
- Failure to inform potential and actual beneficiaries of their rights and entitlements; and,
- Weaknesses in advocacy for national commitments to agreed policy and program commitments, most often because of a failure to target the right key decision makers or a lack of agility in addressing a changing political environment.

#### 3.2.4 Summary and Implications: Objectives Achievement

The evaluations reviewed report a strongly positive, and moderately improved level of objectives achievement by UNICEF supported programs when compared with results in the 2012. The types of positive results reported focus most often on improvements in capacities at the individual (skills, attitudes and behaviours), community (strengthened local

structures for assessing needs and taking action) and national (strengthened national policies, plans and programs) levels.

Not surprisingly, the most frequently cited positive factor contributing to objectives achievement in the evaluations reviewed in 2016 concerns support to well designed training initiatives.

However, there remains scope to raise the proportion of UNICEF supported programs reporting positive results for objectives achievement by addressing important negative contributing factors by:

- Addressing weaknesses in program/project design by improving the clarity and strength of the theoretical linkages from program activities to outputs and outcomes and strengthening program monitoring;
- Investing in the development of baseline data and collecting ongoing information on target group needs; and,
- Including specific measures to better target vulnerable populations including efforts to identify who these populations are and where they are located and program measures to ensure their participation.

## 3.3 CROSS-CUTTING THEMES

### 3.3.1 Coverage

The review used three sub-criteria to address the overall issues related to cross cutting-themes.

- 
- 3.1 Programs effectively address gender equality
  - 3.2 Changes are environmentally sustainable
  - 3.3 Programs effectively address the issue of equity
- 

Coverage of these criteria is not as uniformly strong in the evaluations reviewed as it was for the other criteria, mainly because of the weak

coverage of sub-criterion 3.2 which assesses the extent that changes achieved with UNICEF support are environmentally sustainable (Figure 7). This was addressed in only 12 evaluations and, as a result of this weak level of coverage, findings are not assessed for sub-criterion 3.2.

On the other hand, there are two important areas of improved strength of coverage when comparing the 2016 and 2012 DERs. First, the level of coverage of sub-criterion 3.1 on how UNICEF programming effectively addresses gender equality is much stronger in 2016 than it was in 2012. In the prior review, this sub-criterion was covered only to the moderate level, with 42 of 62 evaluations addressing gender equality. For the 2016 review, 58 of the 66 evaluations in the sample adequately addressed sub-criterion 3.1. It is clear that UNICEF evaluations were much more engaged with the issues of gender equality in the 2012 to 2015 period, than they were from 2008 to 2011.

Sub-criterion 3.3 on how UNICEF programming addresses issues of equity for children and adolescents is new to the 2016 DER. It was added following the inception phase of the review to

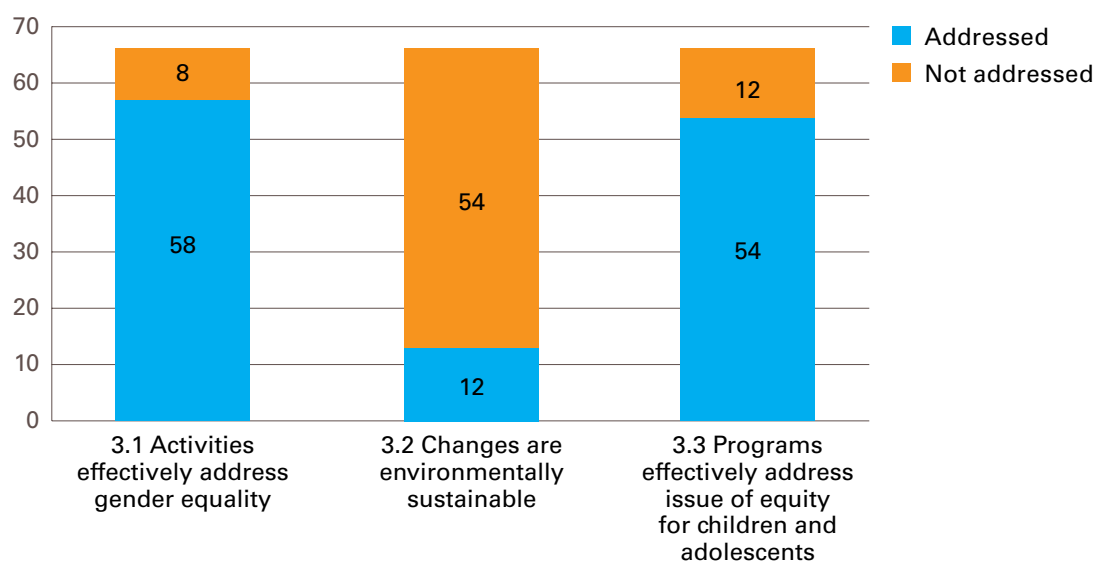
reflect the importance of promoting equity as a key cross cutting strategy for UNICEF during the period under review. Coverage of this new sub-criterion was strong, which indicates that evaluations undertaken by UNICEF in the 2012 to 2015 period were attentive to the critical importance of equity as a cross-cutting theme in UNICEF programs.

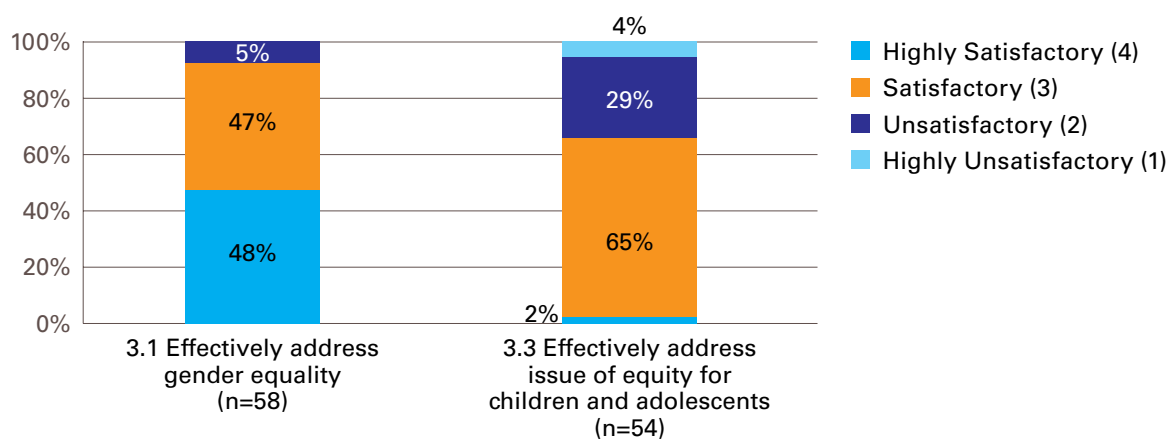
### 3.3.2 Key Findings in Gender Equality

#### Key Findings on Gender Equality

The findings for sub-criterion 3.1 on effectively addressing gender equality were as mixed for the 2016 DER review as they were in 2012. In fact, the proportion of evaluations with satisfactory or highly satisfactory findings on gender equality improved by just one percent between the 2012 and 2016 reviews (47 percent of reviewed evaluations in 2012 and 48 percent in 2016). There was some improvement however, in that the proportion of reviewed evaluations with highly unsatisfactory findings for addressing gender equality was much lower (five percent) in the evaluations reviewed in 2016, than in 2012 (14 percent).

**FIGURE 7** Number of Evaluations Addressing Sub-Criteria for Cross-Cutting Themes



**FIGURE 8 Findings on Effectiveness in Addressing Gender Equality and Promoting Equity**

It is interesting that the current review indicates that findings for effectiveness in addressing gender equality are somewhat more likely to be positive for evaluations of programs in LICs (56 percent) and LMICs (52 percent) than for UMICs (35 percent). This may reflect UNICEF's relatively greater importance as a critical partner to lower income countries and, hence, a greater impact on national policies and practices in gender equality.

The 2015 UNICEF evaluation of UNICEF's strategies and programs to protect children against violence did find some positive results around the issue of gender equality. It noted that attention to gender issues was stronger when UNICEF engaged in efforts to protect girls from violence.<sup>30</sup>

The evaluations reviewed for the 2016 review do cite important successes in addressing gender equality for UNICEF programs in the 2012 to 2015 period. A few examples include:

- Ensuring that women were able to improve their capacities to manage family life and village affairs through training to allow

them to participate fully in village water management committees (Burkina Faso and Burundi);

- Empowerment of women as heads of households in agricultural initiatives and non-traditional employment (El Salvador);
- Providing women with access to direct wage employment and protection from loss of income with an associated improvement in child welfare (Botswana);
- Participation by women in awareness raising activities to engage in local political action as measured by participation and voting in the election of local committees (Sudan);
- Increased participation and access by women and girls accessing health services and taking advantage of micronutrient programs (Ethiopia);
- Increased participation and educational attainment by girls (South Africa);
- Improvement in the knowledge and attitudes of participating men and boys toward issues of gender-based violence (South Africa);

<sup>30</sup> Protecting Children from Violence – A Comprehensive Evaluation of UNICEF's Strategies and Program Performance, 2015. P. xii.

“Training for women in the villages of the province of Ganzourgou strengthened the capacity of the latter in the management of their family life. Women were also well represented in the various bodies established. At the core community contact body, women were three of five members and represented 50% of community contacts. In grassroots associations, the project promoted the representation of women in positions of responsibility. These capacity enhancements of women contribute to consolidating their status in the community and have brought about changes in the relations between men and women. Women strengthened with new behaviours, gain more respect from men.”

*Evaluation finale du projet: Approvisionnement en Eau Potable, Assainissement de base et Hygiène dans les provinces du Ganzourgou et de la Gnagna, Burkina Faso – 2007-2010. 2012 p.72.*

- Improvements in awareness of the rights of girls and women among students and improvements in girls' confidence regarding their chances of succeeding at school (Palestine);
- Changes in attitudes among women and men so that they increasingly withhold support for Female Genital Mutilation/Cutting (FGM/C) (Somalia); and,
- Effectively addressing girls' needs for specific interventions in psychosocial support as a result of the war in Syria (Jordan).

It is important to recognize that when UNICEF supported programs pay adequate attention to addressing issues of gender equality, the reviewed evaluations identify examples of concrete benefits for girls and women, and of increased effectiveness in addressing the rights of children.

### 3.3.3 Contributing Factors in Gender Equality

#### Positive Contributing Factors: Gender Equality

Many of the evaluation reports reviewed identified specific measures aimed at targeting girls and women participants as an important factor in positively addressing gender equality. These measures included: having specific activities for girls and women only (Namibia); targeting

female-headed households (Barbados); and working through organizations representing the interests of girls and women (Mali).

Some evaluations identified the following as positive contributing factors:

- Including women and girls in committees at community level including water and sanitation committees, environmental committees and committees for disaster management;
- Ensuring that projects and programs supported by UNICEF included baselines and ongoing reporting of gender disaggregated information;
- Ensuring that training and communications material was customized to meet the needs of women and girls (and of boys and men) to both enhance capacity and improve understanding of the rights of girls and women.

A few evaluations pointed to the following positive factors in effective support to gender equality:

- Provision of institutional support to organizations representing the interests of girls and women; and,
- Consistent and continuous messaging aimed at improving knowledge and changing attitudes among girls and women and boys and men, including on gender-based violence.

**BOX 11****Assessing the Impact of Child Support Grants on Girls Participation in South Africa**

“There is evidence from gender disaggregated data that the CSG increases the duration of attendance by girls (p.50). There are large differences in impact when we disaggregate by child sex. Early CSG enrolment has positive impacts on many dimensions of girls schooling and learning. Early receipt of the CSG reduces delayed school entry of girls by 26.5 percent. Girls who were enrolled early obtain higher marks on tests of mathematical ability and reading.”

*The South African Child Support Grant Impact Assessment. 2012. P. 66.*

**Negative Contributing Factors: Gender Equality**

In terms of factors with a negative influence on achieving success in addressing gender equality in UNICEF supported programs, by far the most frequently noted in the 2016 review was the simple absence of a gender perspective or the non-application of a gender analysis to the project or program design. Lack of gender analysis was cited in most of the reviewed evaluations which classified findings on gender equality as unsatisfactory or highly unsatisfactory.

A few evaluations indicated that the following represented negative factors for achieving success in addressing gender equality in UNICEF supported programs:

- Absence of gender disaggregated data in baselines and in program monitoring information;
- Time frames which were too short to effect changes in attitudes to gender combined with overly ambitious objectives; and,

- Inappropriate employment of male community mobilizers who failed to encourage participation by girls and women.

In one evaluation report, the authors noted that the overall success of the project may, in fact, have led to reinforcing the gender divide by adding to the workload of women and children (Pakistan).

In summary, the most important factors both contributing to and detracting from success in effectively addressing gender equality in UNICEF programming are concerned with the quality of gender analysis and the inclusion of specific gender equality objectives (and associated design measures) in program designs. Where these are present, they represent a strong positive factor for success; where they are missing, evaluations are much more likely to report negative findings on UNICEF's contribution to addressing gender equality.

**BOX 12****Absence of a Gender Perspective in Program Planning and Monitoring in Montenegro**

“As far as gender equality is concerned, the program logical framework and theory of change does not include any reference to gender issues. The formulation of target groups and project indicators is gender blind. Progress reports do not report gender disaggregated data nor do they contain specific discussions of any gender issues raised during the implementation of the project. The indicators which could be retrieved based on the entries recorded in the national data base on child protection, introduced with the support of the project are not gender disaggregated either.”

*Final Evaluation of the Child Care System Reform in Montenegro. 2014. P. 78.*

### 3.3.4 Summary and Implications: Gender Equality

While UNICEF evaluations have greatly improved their attention to the issue of gender equality since 2012, the reported results are still very challenging with just about half of the evaluations reviewed (48 percent) reporting positive findings (satisfactory or highly satisfactory).

The persistence of these challenging evaluation results needs to be placed in perspective regarding developments in gender equality programming at UNICEF in recent years. These include:

- The establishment of UNICEF's Gender Policy in 2010;
- The development of a three-year Strategic Priority Action Plan (SPAP) on gender equality to cover the period 2010 to 2012;
- The extension of the SPAP to cover 2013;
- The end-of-cycle review to assess the implementation of the SPAP, carried out in 2013 and,
- The development of the UNICEF Gender Action Plan (GAP) (2014-2017) in April 2014, submitted to the June meeting of the Executive Board.

In summarizing the results of the end-of-cycle review, the Gender Action Plan noted:

*"The results of the review indicated that while institutional systems and processes in support of gender equality have significantly been enhanced, a stronger focus is needed on gender programming and results."<sup>31</sup>*

Part of the stronger focus on gender programming and results promised by the Gender Action Plan is found in the establishment of four corporate priorities on gender:

- 1 Promoting gender-responsive adolescent health;
- 2 Advancing girls' secondary education;
- 3 Ending child marriage; and,
- 4 Addressing gender-based violence in emergencies.

Of course, the evaluations sampled for this review, published before the end of 2015, are not expected to reflect the impact of the GAP on UNICEF's effectiveness in addressing gender equality.

The findings of the 2016 review clearly support the need, articulated in the 2014 GAP, for UNICEF to take measures to improve results in gender equality. More specifically, they point to the need for UNICEF supported programs to:

- Adopt a gender perspective when assessing needs of target group members;
- Systematically apply gender analysis during project and program design; and,
- Ensure that projects and programs are designed to include specific measures aimed at targeting (and privileging) participation by girls and women.

The continued absence of a gender perspective in the design of UNICEF supported projects and programs combines with a failure to incorporate gender-specific elements and targets into programs in a way which continues to hamper effectiveness in promoting gender equality.

### 3.3.5 Key Findings on Addressing Equity for Children and Adolescents

Sub-criterion 3.3 on UNICEF programs' effectiveness in contributing to equity for children and adolescents (Figure 8) is new to the 2013 review - so results cannot be compared to the earlier study. However, it is noteworthy that results for this new sub-criterion are positive

<sup>31</sup> United Nations Children's Fund, UNICEF Gender Action Plan, April 2014, p. 4.

with 67 percent of the evaluations reviewed reporting findings which are satisfactory or highly satisfactory.

The evaluations identified both specific targeted group members and different areas of intervention as evidence of effectively addressing equity through UNICEF supported programs. Some examples include:

- A project focus on identifying and overcoming barriers for participation by adolescents and youth (an underserved population) through improved skills and self confidence (Egypt);
- Promoting the right to play and right to education in hard to reach communities (Zambia);
- Improving access to quality education for children from vulnerable groups (Benin) or in hard to reach, high-poverty rural areas (Ghana);
- Improving access to clean water at schools with benefits for participation and education attainment by adolescent girls (Burundi);
- Promoting higher use of Insecticide Treated Bed Nets (ITNs) among the poorest, most rural socio-economic groups (Ghana);
- Providing foster care to children (and among families) of ethnic minorities normally excluded from the service as it is provided to majority populations (Bulgaria);

- Focusing income security programming on poor households most vulnerable to poverty (Botswana and Barbados-St. Lucia);
- Focusing support on hard to reach child workers working in urban settings (Bangladesh);
- Targeting educational assistance to schools where the majority of students come from marginalized and tribal communities (India); and,
- Focusing the program on high risk groups including migrant workers, construction workers, brick kiln workers, nomads and residents of urban slums (India).

While not exhaustive, the list of examples illustrates how many UNICEF supported programs go beyond a general target of children and women to target underserved and marginalized populations including: those who are geographically isolated (and so denied services), are members of ethnic minorities, are among the poorest segments of the population or have specific barriers to participation such as those faced by adolescent girls in schools without adequate water supplies.

### 3.3.6 Contributing Factors on Addressing Equity for Children and Adolescents

The most frequently noted positive factor contributing to success in addressing equity for children and adolescents, cited in many evaluations, was an explicit targeting strategy which

#### BOX 10

#### Using Radio to Reach Poor Rural Children in Zambia

“The strategy applied by the Interactive Radio Instruction (IRI) project was relevant in that the project was implemented in the rural areas with the lowest coverage rate of participation in early childhood education (ECE) for the target age group. Most of the project sites within the four districts were located in hard-to-reach area where children did not have access to ECE. Even in sites where ECE services existed, vulnerable children could not afford to access them. In this regard, implementation of the IRI project in the four districts has enhanced equality of access to ECE and has promoted children’s rights such as the right to play and the right to education.”

*Evaluation of the Early Childhood Education Interactive Radio Instruction Pilot Program in the Eastern Province of Zambia. 2015. P. 16.*

focused the benefits on marginalized and underserved individuals and communities. This most often involved defining criteria for participation which either focused exclusively on members of the marginalized group or provided them distinct advantages in the process of qualifying to participate.

A few evaluation reports highlighted the positive effect of the following factors on success in promoting equity:

- Use of appropriate technologies to meet the needs of populations in hard to reach areas where conventional medical testing equipment was not available;
- Engaging members of hard to reach and marginalized groups in project design; and,
- Use of a life cycle approach in project design, so that interventions were addressing the right of children to survival, growth and development.

#### **Negative Contributing Factors: Equity for Children and Adolescents**

Many of the evaluations reporting negative findings on effectively addressing equity for children and women highlighted the absence of any clear focus on the equity aspects of the program and/or the needs of vulnerable and marginalized populations. This was sometimes linked to the absence of any analysis of bottlenecks and barriers to access for vulnerable populations.

Some evaluations pointed to lack of information on the conditions and needs of vulnerable group members and/or poor monitoring of the differential effect of programs on specific groups.

A few evaluations also cited the negative effect on equity of the following factors:

- A tendency to focus on the easiest to reach and most readily identified children and women for program support;
- Poor infrastructure in schools which, if not addressed by the program, discourages access for special needs children; and,

- Inappropriate technology which, while it may be lower in cost, inhibits inclusiveness and discourages participation by disabled persons.

As with sub-criterion 3.1 on gender equality, it seems that the most important factors for improving success in addressing equity for children and women in UNICEF supported programs are concerned with the need to first analyze the barriers to equity and then to develop explicit strategies, goals and objectives. Doing so helps to ensure that vulnerable group members participate and benefit from programming.

#### **3.3.7 Summary and Implications: Addressing Equity**

The overall pattern of findings on UNICEF's success in addressing equity for children and adolescents is largely positive but there remains considerable scope for improvement (over 30 percent of evaluations addressing equity report negative findings).

A review of the positive and negative factors which either impede or enhance UNICEF's success in promoting equity suggests that positive results can be achieved by more programs by:

- Better defining and identifying who are the most disadvantaged, where they are located and how they access services (including explicit and systematic identification of barriers);
- Adopting an explicit strategy for targeting programs to focus the benefits on the most disadvantaged which often involves defining criteria for participation which confers real advantages on target group members; and,
- Investing in improved information on the location, conditions and needs of vulnerable group members which can be linked to monitoring the effect of programs on specific groups.

## 3.4 SUSTAINABILITY

### 3.4.1 Coverage

The review addressed the overall question of sustainability through three sub-criteria which can be expressed as positive statements:

- 4.1** Benefits are likely to continue after project or program completion.
- 4.2** Projects and programs are reported as sustainable in institutional and/or community capacity.
- 4.3** Programming contributes to strengthening the enabling environment for development.

Coverage for all three sub-criterion relating to sustainability was rated as strong in the 2016 review (Figure 9). This was an improvement on the coverage in the 2012 review, with coverage of sub-criterion 4.3 on strengthening the enabling environment rated then as only moderate (40 of 59 evaluation reports). This improvement in coverage may be a result of UNICEF programming which is more focused on upstream work

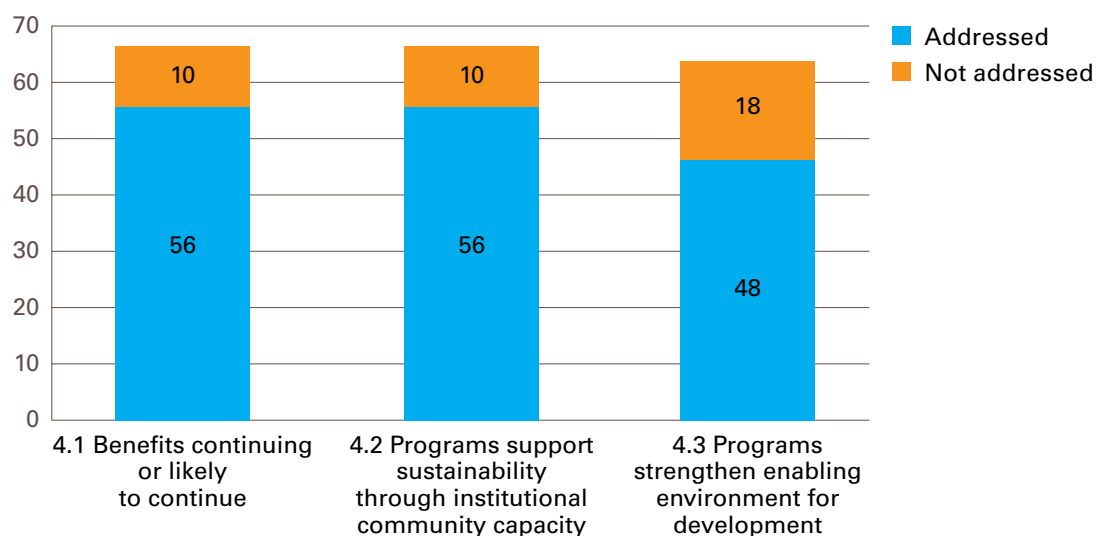
in advocacy and policy engagement in the 2012-2015 period when compared to the 2008 to 2011 period covered by the previous review.

### 3.4.2 Key Findings on Sustainability

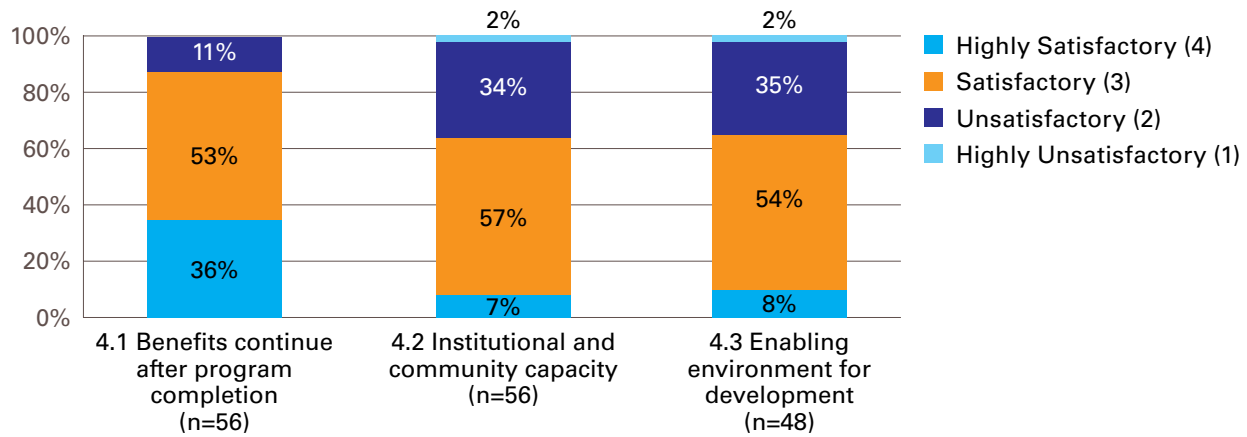
Overall, the evaluation reports reviewed for this DER of UNICEF report highly mixed results on evaluation sub-criteria relating to sustainability. Sub-criterion 4.1 on the likelihood of benefits continuing in the future, after program completion received positive findings (satisfactory or highly satisfactory) in only 36 percent of the evaluations which addressed it. On the other hand, findings were largely positive for sub-criterion 4.2 on the capacity of institutions and communities to sustain program activities (64 percent). Similarly, 62 percent of the evaluations addressing sub-criterion 4.3 reported positive findings regarding the effect of UNICEF programming on the enabling environment for development; an important consideration for sustainability.

When comparing results of the 2016 and the 2012 reviews under the criteria of sustainability, it is clear that some aspects are improved while others have apparently weakened. In

**FIGURE 9** Number of Evaluations Addressing Sub-Criteria for Sustainability



**FIGURE 10 Findings on Sustainability**



2012 (covering evaluations from 2009 to 2011) just over half (51 percent) of the evaluations reviewed reported positive findings on sub-criterion 4.1 on the likelihood of benefits continuing after program completion. For the 2016 review (covering evaluations from 2012 to 2015) this proportion had dropped to just 36 percent.

The proportion of evaluations reporting positive results for sub-criterion 4.3 on the enabling environment for development also declined from 71 percent in the 2012 review to 63 percent in 2016. On the other hand, the proportion of evaluations reporting positive findings for sub-criterion 4.2, on improved institutional and community capacity to sustain programs, improved to 64 percent in the 2016 review, from only 48 percent in 2012.

The review team noted that the evaluations covered in the 2016 review were much more focused on issues of financial sustainability than those covered in the 2012 review. This heightened attention to issues of sustainability may have strongly influenced the decline in the number of evaluations reporting positive results for sub-criterion 4.1 on the continuation of program benefits. If evaluations conducted in the period 2008 to 2011 did not pay the same

level of attention to issues of financial sustainability, they may have had a bias to more positive findings than those undertaken from 2012 to 2015.

As a result, it is not clear if the decline in positive findings for sub-criterion 4.1 on the likelihood of benefits continuing results from an actual shift in the sustainability of programs or a much higher level of concern in UNICEF evaluations on the financial sustainability of program results.

Problems in achieving sustainability were also noted in global evaluations undertaken by UNICEF during the review period. The 2014 evaluation of upstream work in education and gender equality noted that “like most donors, UNICEF has struggled to ensure the sustainability of capacity building efforts”<sup>32</sup>

From a country classification perspective; evaluations of programs in UMICs reported positive findings for sub-criterion 4.1 on the likelihood of benefits continuing (47 percent) and for sub-criterion 4.2 on the sustainability of institutional and community capacity (79 percent), more often than evaluations of programs in LICs or LMICs.

<sup>32</sup> UNICEF’s Upstream Work in Basic Education and Gender Equality – 2003 to 2012. (2014) P.13.

There are many dimensions of sustainability, across all three sub-criteria which have been strengthened by the programs addressed in the evaluations reviewed in 2016. These include:

- Training of community members who are capable of maintaining water infrastructure and using new methods to improve crop yields (Fiji);
- A modernized policy and legislative framework which provides for more sustainable and coordinated service provision to excluded and vulnerable children and their families (Montenegro);
- Implementation of community-based strategies and social protection initiatives for children and women in vulnerable situations which, in turn, contributes to a dynamic favourable to sustaining achievements at community level (Niger);
- Integration of a gender-based violence component into the ongoing work programs of implementing partners (South Africa);
- Development of a framework and national commitment to integrate the Child Friendly Schools (CFS) model into the entire educational system (Moldova);
- Governmental allocation of funding for institution building and the development of services based on continuing institutional reforms (Montenegro);
- Development of a good level of understanding of operational difficulties and possible solutions, based on testing by the implementing partners (Mozambique);
- A strengthened institutional framework for the protection of the child through the creation of a specific departmental unit in the Civil Registry Office and the training of the judiciary dealing with minors, as well as special police brigades to deal with minors (Niger); and,

- Provision of technical support to government institutions which will assist them in future in developing and successfully enforcing relevant juvenile justice legislation in line with international standards (Bosnia).

### 3.4.3 Contributing Factors

#### Positive Contributing Factors: Sustainability

The evaluations reviewed for the 2016 DER identified a number of factors which, when incorporated into UNICEF supported programs, tended to strengthen sustainability across one or more of the three sub-criteria assessed.

Most evaluation reports emphasized the importance of active involvement of community members and community based committees, management structures or pre-existing community based organizations as a strong positive determinant of the sustainability of results.

Many evaluation reports pointed to the capacity of national and local government institutions as an important factor positively effecting sustainability. Often this was linked in the evaluation report to effective capacity building efforts supported by UNICEF (Bosnia, Mozambique).

Some evaluation reports noted as positive factors for sustainability:

- Integration of UNICEF supported program activities into a wider framework of national policies and programs (including sector strategies and budgets);
- Engagement by UNICEF in efforts to strengthen the institutional and policy framework in which the program operates, including effective work on developing national strategies, policies, programs and technical guidelines;
- National and local government buy-in and a sense of ownership by national and local authorities;
- Effective partnerships among UNICEF, national government agencies, local service providers and NGO implementing partners;

“The continued existence of Community Development Committees (CDCs), cooperatives and other community-based organizations in an important indicator of the continued benefits of the project. This means that skills and knowledge are being used and contributing to the resilience of communities. Ensuring that people, in particular, state officials, have the capacity to fulfill their roles is essential to ensure that additional work will be conducted and support requested to address ongoing and emerging needs.

Through discussions with CDCs and community mobilizers, communities have taken ownership of project initiatives. Their continued efforts are demonstrated through their efforts to protect life during hurricane Pam.”

*Community Resilience and Coping with Climate Change in Vanuatu. 2015. P.44.*

- An emphasis in program design and implementation on the need for inter-sectoral work, including for example health and food and nutrition (El Salvador), education and child protection (Turkey), and education and science as well as child protection (Zambia); and,
- Well-designed training and supportive supervision elements which helped to provide administrators and service providers with skills and capacities to sustain program activities.

#### **Negative Contributing Factors: Sustainability**

By far the most frequently cited factor which negatively influences the sustainability of UNICEF programming is the failure to either generate sufficient finances locally or secure adequate budgets from national authorities to continue program activities. Most evaluations with negative findings on sustainability pointed to weak finances as a critical factor.

Many evaluation reports pointed to weak institutional capacity, especially at the local level and including, for example, weak water management committees (Burundi). This weakness in capacity was also linked to poorly designed training programs or simple failure to address capacity problems.

Some evaluations identified exit strategies that were either weak or entirely missing as a key factor limiting the sustainability of the results of UNICEF programming.

A few evaluations noted the following as factors limiting sustainability:

- Compressed project or program time frames which meant resources were withdrawn before sustainability could be achieved;
- Exclusionary cultural factors and negative attitudes toward women which limited the sustainability of gains achieved;
- Lack of effective community participation; and,
- High turnover of staff in key participating government bodies and institutions.

In summary, the reviewed evaluations indicate that programs supported by UNICEF (and their associated results) are most sustainable when they are grounded at local level through active community engagement, when adequate sources of financing are identified, and when UNICEF has been engaged and effective in the institutional and policy framework in which the program operates.

**BOX 14****Lack of Finance and Government Buy-In Impede Sustainability in Tanzania**

"The implementation of a product such as electronic Maternal and Newborn Health (eMNH) and Nutrition (eNutrition) requires a substantial resource investment both in terms of human resources for technical support and in terms of financial resources, making it difficult to conceive of a model where the application could go to any scale or to foresee cost sharing with government. (p.8)." "Consideration should be given to the extent projects have sustained buy-in and ownership from national stakeholders and ensuring a more structured process is initiated as the basis for feasibility assessment and building the evidence base for demonstrating the value of pilot projects on a broader scale." (p.9).

*Evaluation of Mobile Decision Support Applications in Tanzania. 2013. P. 8-9.*

### 3.4.4 Summary and Implications: Sustainability

UNICEF's evaluations have improved their coverage of sustainability in the period from 2012 to 2015. However, the findings they report remain challenging, although there has been some improvement with regard to the sustainability of efforts to support the capacity of communities and institutions to sustain the results of UNICEF supported programs.

UNICEF evaluations have also focused much more closely on the financial sustainability of supported projects and programs, particularly their ability to generate finances internal or to attract continued support from host governments and external donors. The evaluations reviewed provide evidence that UNICEF could make gains in the sustainability of results by:

- Ensuring during program design and development that supported initiatives directly address the question of financial sustainability by:
  - Including realistic measures to generate income locally;
  - Secure commitments from national/regional and local authorities for budgeted resources to continue program initiatives after the end of UNICEF support;
  - Identify sources of external financing where those are needed to supplement host government budgets; and,

- Ensure that supported programs include a realistic assessment of the institutional capacity of implementing partners and adequate measures to address any identified inadequacies.

## 3.5 EFFICIENCY

### 3.5.1 Coverage

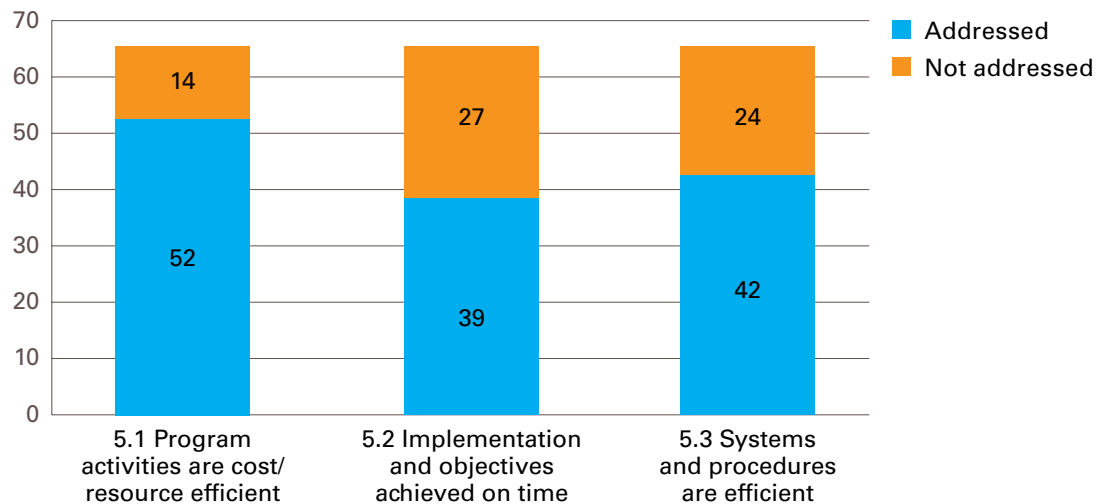
The review addresses the overall criterion of efficiency through three more focused sub-criteria:

- 
- 5.1** Program activities are evaluated as cost/resource efficient.
  - 5.2** Implementation and objectives achieved on time.
  - 5.3** Systems and procedures for project/program implementation and follow up are efficient.
- 

The pattern of coverage for the three sub-criteria relating to efficiency is almost identical for the 2016 review as it was in 2012. In both cases, sub-criterion 5.1 on the cost and resource efficiency was rated strong in coverage. Similarly, for both the 2016 and 2012 reviews coverage of sub-criterion 5.2 on timely achievement of objectives and 5.3 on the efficiency of systems and procedures for program implementation was rated only moderate.

**FIGURE 11**

**Number of Evaluations Addressing Sub-Criteria for Efficiency**



For sub-criterion 5.1 on cost/resource efficiency, the review team did not require a formal cost effectiveness assessment to indicate that the criterion had been addressed in an evaluation. Rather, in both 2016 and 2012, the team considered this criterion adequately addressed if the evaluation report assessed (positively or negatively) the costs of the supported program in relation to the scale of its contribution to results.

It is worth noting that the review team indicated there may be a negative bias in the findings reported for sub-criteria 5.2 (on timeliness) and 5.3 (on the efficiency of administrative systems and procedures). These sub-criteria are not often formally included in evaluation terms of reference. As a result, they may be much more likely to be addressed in evaluations where the findings are negative than would occur with a positive result. This perception is supported somewhat by the significant number of evaluations which do not address sub criterion 5.2 and 5.3.

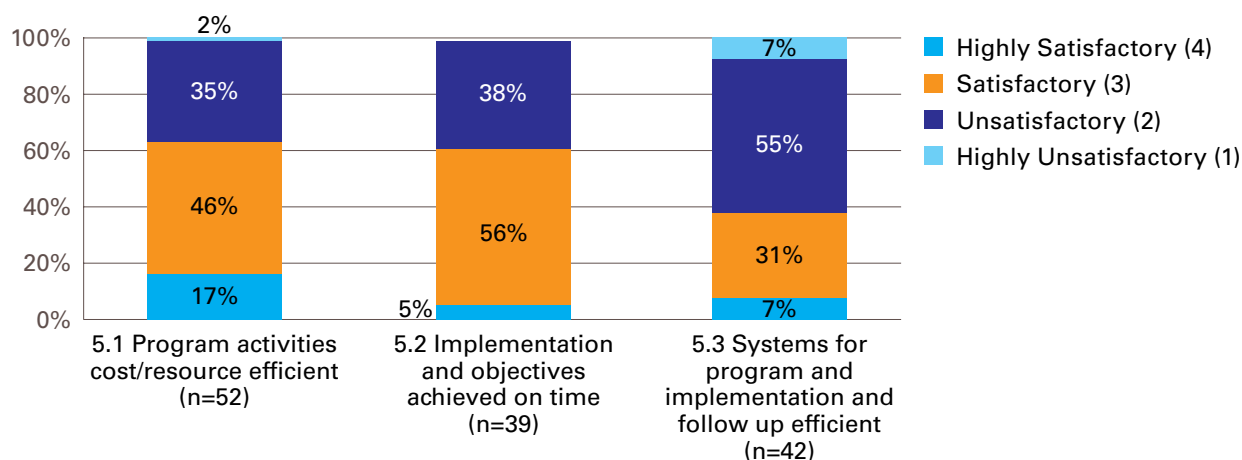
### 3.5.2 Key Findings

The reviewed evaluations present mixed results across the three different sub-criteria relating to efficiency and each should be considered separately.

For sub-criterion 5.1 on the cost and resource efficiency of UNICEF supported programs, the findings are broadly positive. Fully 63 percent of the evaluations addressing sub-criterion 5.1 report that UNICEF support activities were cost efficient. Remarkably, 17 percent of the evaluations reviewed found program performance to be highly satisfactory with regard to cost efficiency.

It is worth noting that the results for the 2016 and 2012 reviews were very similar on this point with 62 percent of evaluations reporting positive findings in the 2012 review versus 63 percent in 2016.

From a country classification perspective, it is interesting that evaluations of programs in LICs were more likely to report positive (satisfactory or highly satisfactory) findings on cost efficiency (75 percent) than evaluations of programs in LMICs (59 percent) and UMICs (57 percent).

**FIGURE 12** Findings on Efficiency

The reviewed evaluations were also mainly positive on the timely achievement of program objectives, with 61 percent of evaluations reporting satisfactory or highly satisfactory results for sub-criterion 5.2 in the 2016 review. This is a significant improvement on the 2012 review, which found that only 51 percent of reviewed evaluations reported positive findings on timeliness.

However, findings for sub-criterion 5.3 on the efficiency of systems for program implementation (including systems for engaging staff, procuring inputs, disbursing payments, and program logistical arrangements) and follow-up were negative, with only 38 percent of evaluations reporting findings as the satisfactory or highly satisfactory level. This negative pattern of findings was very similar for the 2012 review, when the proportion of evaluation reports with positive findings for sub-criterion 5.3 was only 37 percent.

Those evaluation reports which indicated that UNICEF supported programs were efficient highlighted different aspects of efficiency:

- An ability to reach more beneficiaries or exceed the planned level of outputs within the original program budget (Egypt, Burkina Faso, Jordan);
- The ability to manage limited resources frugally and keep management structures and associated administrative costs to a minimum (Haiti, Macedonia, India);
- The achievement of low unit costs (Bulgaria, Uzbekistan);
- Avoiding duplication and overlap in activities across agencies (Fiji);
- Effectively making use of volunteers and interns to produce knowledge products (India);
- Overcoming time-related obstacles and operational delays so that delays in producing outputs and contributing to outcomes were limited (Egypt, Somalia, Uzbekistan);
- Strong technical support by UNICEF which helped to overcome challenges in procuring and distributing commodities (Tanzania); and,
- Flexibility on UNICEF's part so that program plans could be quickly adjusted in response to changing circumstances and contexts (Egypt, Colombia, DRC).

“The Internship Program (IP) may be considered efficient because the cost of implementing the program has remained consistent over the years. Stakeholders, including host organizations and interns contribute to the resource base of the program. The interns, particularly the international interns, fund their own travel costs. The host organizations receive only modest support for the cost of housing and field work. The case studies as a knowledge management product are cost effective considering the short time frames and high level of skills available to produce them by a team of young professionals using original field data”

*Evaluation of Knowledge Community on Children in India Internship Program. 2012. P.6.*

### 3.5.3 Contributing Factors

#### Positive Contributing Factors: Efficiency

The range of factors making a positive contribution to efficiency was very diverse with no single group of factors noted in a large number of evaluations. As a result, the following factors contributing positively to efficiency were each noted in a few evaluations:

- Efficient UNICEF funding and disbursement practise which sometimes made it possible to avoid delays in disbursing funds through centralized national systems;
- Lean management structures with associated low overhead costs for UNICEF supported programs;
- Careful planning of UNICEF supported programs in order to avoid duplication and overlap, especially among different supporting agencies;
- Use of centralized purchasing arrangements to achieve lower input costs;
- Use of low cost technologies to deliver educational services;
- Implementing programs through local NGOs, civil society organizations and implementing partners rather than through national or international organizations;

- Helping to strengthen referral systems in health support programs so that procedures are carried out at local level and at much lower cost; and,
- Overall, strong capacity for project management within UNICEF country offices.

#### Negative Contributing Factors: Efficiency

The range of factors making a negative contribution to efficiency was somewhat more concentrated with some evaluation reports pointing to:

- Problems and delays in the disbursement of funds from UNICEF to implementing partners;
- Human resource challenges, including rapid turnover of staff, delays in hiring staff or under-staffing in UNICEF country offices;
- Lack of sufficient data on both costs and outcomes to allow effective management or monitoring of program efficiency; and,
- Over-ambitious project objectives, especially concerning delivery of outputs (such as training courses, capacity development interventions, equipment purchases, etc.) within very limited time frames.

**BOX 16****Staff Turnover Impedes Efficiency in the Sudan**

“The staffing problem is exacerbated by the high rate of staff turnover due to low salaries. This problem is multi-dimensional: besides sustainability it also effects effectiveness and efficiency, since the staff members who leave are normally those who are well qualified and have received substantial training. Moreover, it hinders interaction and trust building with partners and communities.”

*Evaluation: UNICEF Assisted Water, Sanitation and Hygiene Programs in Sudan (2002 to 2010). 2012. P. 86.*

### 3.5.4 Summary and Implications: Efficiency

UNICEF supported programs continue to be found efficient in their use of financial and non-financial resources to achieve results. In addition, the evaluations reviewed showed an improvement in the percentage of evaluation reports with positive findings relating to the timeliness of program implementation. On the other hand, evaluations which address the efficiency of UNICEF administrative systems and procedures continue to report mainly negative findings (although there is some reason to believe that the percentage may be overstated). Improving the efficiency of UNICEF supported programs and projects will require:

- Continuing to emphasize lean program management structures and associated low overhead costs;
- Careful planning to avoid duplication and overlap when support is provided in the same sector by different external agencies;
- Strengthening monitoring of both costs and outcomes to allow for effective management of program efficiency; and,
- Addressing human resource challenges including rapid turnover of UNICEF staff in country offices and delays in hiring needed staff.



## CHAPTER 4

# EVALUATION AND MONITORING FOR IMPROVED EFFECTIVENESS

## CHAPTER 4

This chapter deals with the results of the 2016 DER as they deal specifically with UNICEF's effectiveness in using evaluation and monitoring systems and processes, including Results-Based Management (RBM) to improve development effectiveness. The separate chapter is used as a means of highlighting the implications of the review for both evaluation and monitoring at UNICEF since identifying potential improvements in both areas is an important objective of the study.

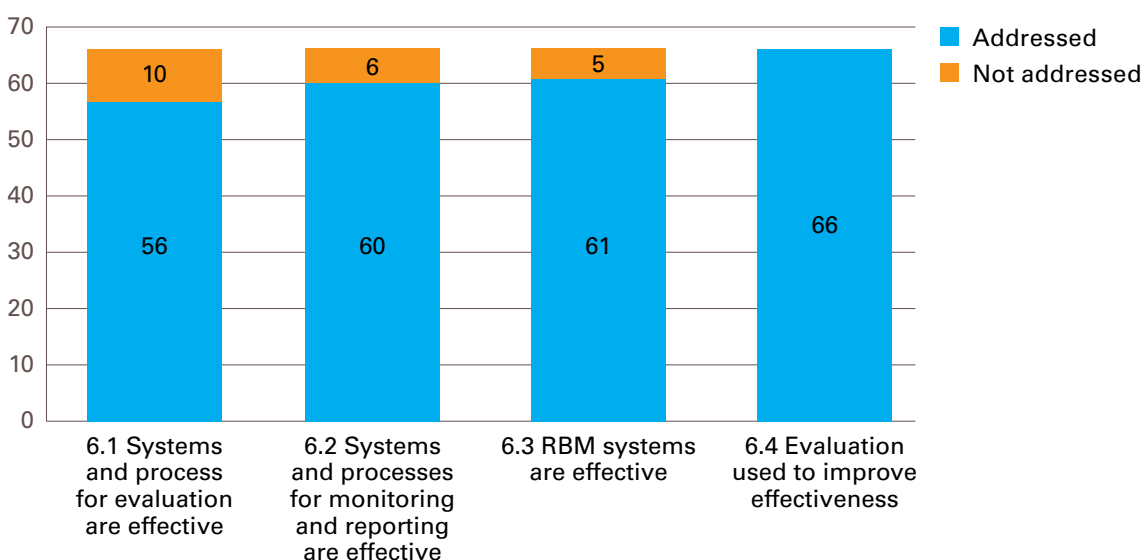
The four sub-criteria used to address the use of evaluation and monitoring include:

- 6.1** Systems and processes for evaluation are effective
- 6.2** Systems and processes for monitoring are effective
- 6.3** Results-based management systems are effective
- 6.4** Evaluation results used to improve humanitarian and development effectiveness

### 4.1 COVERAGE

Coverage of all four sub-criteria relating to the use of evaluation and monitoring to improve effectiveness was rated as strong for the 2016 review (Figure 13). This was a significant improvement over the 2012 review results, in which sub-criterion 6.1 on the effectiveness of systems and process for evaluation was covered only to the moderate level (Table 2), and coverage of sub-criterion 6.3 on the effectiveness of results-based management (RBM) systems was weak (also Table 2).

**FIGURE 13** Number of Evaluations Addressing Sub-criteria for Evaluation and Monitoring



Evaluations carried out from 2012 to 2015 were much more likely to address the strength and effectiveness of RBM systems than those carried out from 2009 to 20011. It seems that the strong emphasis on results monitoring noted in Section 1.7 (with the advent of the Field Results Group in 2014, the ongoing attention paid to MoRES and the use of the Results Assessment Module since 2013) has led to greater attention paid to RBM in UNICEF evaluations.

The findings for overall criteria six “using evaluation and monitoring to improve the effectiveness of development and humanitarian programming” fall into two different groups when broken down by sub-criteria (Figure 14).

## 4.2 FINDINGS: USE OF EVALUATION

### 4.2.1 Key Findings: Use of Evaluation to Improve Effectiveness

Findings for the sub-criteria relating to evaluation and its uses are strongly positive.

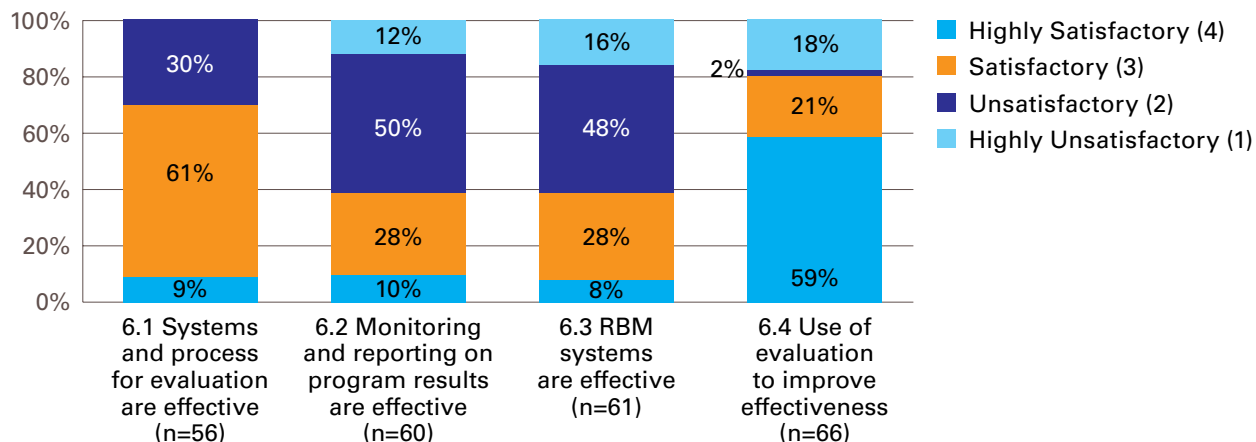
- For sub-criterion 6.1 (on the effectiveness of systems and processes for evaluation), seventy percent of the evaluations addressing the sub-criterion reported findings which were positive (satisfactory or highly satisfactory). This was an improvement over the

results for the same sub-criterion in 2012, when the review found 64 percent of evaluations reported a positive finding;

- For sub-criterion 6.4 on how evaluation results are used to improve development and humanitarian effectiveness (which mainly focuses on the formal management response) results for the 2016 review are very positive with 80 percent of evaluations reporting findings of either satisfactory or highly satisfactory. This was a significant improvement over the result for the 2012 review which reported positive findings for 57 percent of the reviewed evaluations. Even more notable, in the 2016 review, fully 59 percent of reviewed evaluations reported highly satisfactory findings for the use of evaluation to improve effectiveness.

Overall, however, it is important to note that the evaluations reviewed for the 2016 were more likely to report positive results on the use of evaluations to improve development and humanitarian effectiveness than those reviewed in 2012. This is consistent with the team’s observation that evaluations reviewed in 2016 were of generally higher quality and suggests that UNICEF has been effective in improving the quality and utility of evaluations conducted in the past three years.

**FIGURE 14** Findings for Using Evaluation and Monitoring



It is important to note that the 18 percent of evaluation reports which were classified as presenting highly unsatisfactory findings for sub-criterion 6.4 can be accounted for by the 12 reviewed evaluations which did not post a management response at the time of the review. Since eight of those are evaluation reports published in 2015, it is likely that most of them will post a management response in the coming months, which will significantly reduce the proportion of evaluations with negative findings for this sub-criterion. If these 12 evaluation reports are eliminated from the analysis, 98 percent of the remaining 54 evaluations report positive findings for this sub-criterion.

#### 4.2.2 Contributing Factors: Using Evaluation to Improve Effectiveness

Many evaluation reports noted that the existence of a regular schedule of end of project evaluations and mid-term reviews contributed to the effective use of evaluations to improve

effectiveness. They often noted (DRC, Chad, Ghana) that lessons from the evaluation of a previous project or program were used to strengthen the design of the current program.

A few evaluations pointed to each of the following positive factors promoting the use of evaluation results to improve effectiveness;

- Strong baseline data collection allowed for a more robust evaluation with valuable lessons learned;
- A tradition and practice at UNICEF of using external evaluators and external evaluations to support learning;
- A systematic approach to evaluation with a commitment to regular, ongoing evaluations and to covering most programs;
- A commitment to using the results of all types of evaluations (formative, mid-term, end-of program or project) to improve program design.

#### BOX 17

#### Ongoing Evaluation Improves Insecticide Treated Net Program in Ghana

“The quantitative and qualitative data indicate that implementation lessons were learned from the earlier distributions such that many prior mistakes were averted, such as providing a sufficient number of nets to cover all the households. Based on these results, a similar strategy of distribution through house to house visits with hang-up activities was adopted for the other regions in Ghana: including all households to progress towards universal coverage.”

*Evaluation of the LLIN Universal Mass Distribution Campaign in Brong Ahafo, Central and Western Regions of Ghana. 2013. P. 7.*

#### BOX 18

#### Systematic Evaluation Supports Maternal and Child Health Programming in Uzbekistan

“The robust Monitoring and Evaluation system (including supportive supervision) established and applied by the project was effective in reinforcing skills application and in tracking progress at national, sub-national, and facilities levels. The system was established at all levels of the health care system as well as at the project management level within UNICEF. In order to monitor project implementation in the field, the project also supported the Ministry of Health in organizing joint monitoring team visits.” (p. 40).

*Evaluation of Improvement of Mother and Child Health Services Phase II: Uzbekistan. 2015. P.40.*

### Negative Contributing Factors: Evaluation

The most frequent negative comments regarding evaluation and its use in improving effectiveness at UNICEF were focused on those projects and programs which did not seem to be subject to regular evaluations. A few of the reviewed evaluation reports identified the following negative factors relating to evaluation:

- A primary focus on project monitoring with inadequate attention by UNICEF and the implementing partner to the question of evaluation;
- The absence of any commitment to, or mention of, a regular program of evaluation so that there is no culture of evaluation embedded in the program;
- Willingness by UNICEF and other partners to renew programs and enter into a new programming cycle without a prior evaluation; and,
- A focus on reporting outputs and neglecting learning and accountability which reduces the incentive to invest in evaluation.

#### 4.2.3 Summary and Implications: Using Evaluation to Improve Effectiveness

The 2016 review noted continued improvement across the sub-criteria used to assess the extent that UNICEF makes effective use of evaluation as a function to improve effectiveness in achieving results. The trend noted was to continuing improvements in the quality and utility of evaluations. At the same time, the reviewed evaluations identified measures which could help strengthen evaluation processes and improve the use of evaluation for enhance effectiveness. These include:

- Ensuring that all programs are evaluated on a regular basis according to a committed schedule of end-of project evaluations and mid-term reviews;

- Avoiding the practice of renewing programs or designing successor programs without a prior evaluation; and,
- Committing to the more frequent development of project and program baseline at the output and outcome level along with regular monitoring and reporting in order to contribute to improving evaluations.

## 4.3 FINDINGS: MONITORING FOR IMPROVED EFFECTIVENESS

### 4.3.1 Key Findings on Monitoring for Improved Effectiveness

As noted in Figure 14, the findings for sub-criterion 6.2 (on the effectiveness of results monitoring and reporting systems) and 6.3 (on the effectiveness of RBM systems) are largely negative.

For sub-criterion 6.2, the effectiveness of results monitoring and reporting systems, only 38 percent of evaluations reported results which were either satisfactory or highly satisfactory. This is very nearly identical to the proportion of evaluations with positive findings on monitoring and reporting systems found during the 2012 review (39 percent).

For sub-criterion 6.3 on Results-Based Management Systems, only 36 percent of reviewed evaluations reported positive findings. This cannot be readily compared with the results of the 2012 review since coverage of sub-criterion 6.3 was rated weak during that review.

In summary, while evaluations carried out in the 2012 to 2015 time-frame were much more likely to address the effectiveness of RBM systems at UNICEF, they often reported negative findings. The same was true for the findings on the general strength of systems for monitoring and reporting on results (sub-criterion 6.2) with around 40 percent of evaluations

reporting positive findings, a proportion which is largely unchanged from the findings of the 2012 Review.

It is worth noting that the evaluations reviewed would not have reflected the initiatives undertaken by the Field Results Group to strengthen results-based management, especially ongoing efforts in RBM training and the introduction of the country office scorecard and the country office strategy note.

#### 4.3.2 Contributing Factors: Monitoring for Improved Effectiveness

To some extent, of course, factors which contribute positively to ongoing results monitoring and reporting can also be seen as providing a strong basis for evaluation. In that sense, the positive factors noted for sub-criterion 6.2 on the effectiveness of results monitoring and reporting and 6.3 on RBM systems are also relevant to the question of effective evaluation.

Many evaluation reports noted that the early establishment and operationalization of a systematic approach to program monitoring was an essential factor in contributing to effective results monitoring and to RBM.

Some evaluation reports pointed to the following factors contributing positively to the use of monitoring and RBM systems to improve program effectiveness at UNICEF:

- The establishment and tracking of targets at both the output and outcome levels supported by regular reporting on achievements;
- Well defined and documented program theories of change or logic models which can support the definition of indicators and targets for monitoring progress and taking decisions – leading to the production and use of robust results frameworks; and,
- Use of baseline data and regular surveys to augment ongoing collection of program results data and strengthen results-based management.

A few evaluation reports noted each of the following as positive factors promoting effective use of monitoring and reporting systems, including RBM;

- Use of supportive supervision as a method for sustaining data collection and ongoing tracking of program progress by supporting the monitoring skills of staff;
- Use of regular planning and review meetings to emphasize the priority given to monitoring and evaluation;
- Development by UNICEF of effective tools for monitoring program progress and sharing of those tools with implementing partners; and,
- Strong technical assistance to the development of monitoring systems by UNICEF staff.

Not surprisingly, given the predominance of negative findings for sub-criterion 6.2 on monitoring systems and 6.3 on RBM (Figure 14), evaluation reports often identified negative factors detracting from the use of monitoring systems (including RBM) to improve the effectiveness of UNICEF supported programs.

Most evaluation reports pointed to the general absence of a systematic approach to program monitoring or to lack of emphasis on monitoring which further weakened RBM systems. Many also referred to the absence of regular data collection, data missing at periodical intervals, or a general absence of supporting data in the monitoring reports that were being produced by UNICEF supported programs. Data limitations often extended to a general absence of baseline data for use in establishing appropriate targets at both output and outcome levels.

Some evaluations noted the following as negative factors limiting the effectiveness of monitoring systems and RBM:

- Absent or relatively weak theories of change or program logic models;

“In almost all centres sampled in this study, participants indicated that there was no sustained monitoring and evaluation of Interactive Radio Instruction programs. This meant that challenges could not be detected early enough to prevent discontinuation of programs in due course” (p. ix). “The evaluation revealed there was not consistent monitoring and supervision of IRI activities in the four districts, making it difficult to determine whether or not progress was being made in the attainment of objectives.” (p. 36).

*Evaluation of the Early Childhood Education Interactive Radio Instruction Pilot Program in the Eastern Province of Zambia. 2015, P.36.*

- Inappropriate indicators developed at output and outcome levels; and,
- Ad-hoc reporting of results indicators with no integration or consolidation into a regular overview report which can be used for accountability, learning and management.

A few evaluation reports identified the following negative factors, limiting the effective use of monitoring systems to improve program effectiveness:

- Use of inappropriate technology such as web-based data entry where internet access was limited or non-existent;
- Weaknesses in analysis so that trends in output and outcome data are not interpreted in relation to program goals and targets in a credible way; and,
- Poor quality control or quality assurance of data entered into the monitoring systems so that results reported are inconsistent and provide conflicting information over time (which limits the system’s ability to support learning and accountability); and,
- Limited staff capacity among implementing partners for effectively establishing and maintaining results monitoring and reporting systems.

### 4.3.3 Summary and Implications: Use of Monitoring to Improve Effectiveness

In summary, despite the increase in attention given to results monitoring and reporting by UNICEF since 2012, any improvements in the strength of monitoring systems have not been recognizable to evaluations carried out from 2012 to 2015. This may, in part, reflect the retrospective nature of evaluations themselves since the programs under evaluation in those three years most often were established some years earlier.

In some ways, the establishment of the Field Results Group in late 2014 (See Section 1.7 above) can be seen as a direct response to the challenge of strengthening results monitoring and management at UNICEF. The FRG is attempting to strengthen results monitoring and management through its emphasis on: skills development in RBM; improved results reporting for accountability; better monitoring of performance at country office level; and improving modes of work with implementing partners and contractors. However, since the work of the FRG began in 2014, it is unlikely to have had an effect on the programs reviewed from 2012 to 2015.

At the same time, the evaluations examined during this review point to specific steps which can be taken to augment the work of the FRG and

to strengthen results monitoring for improved development and humanitarian program effectiveness at UNICEF. These include:

- Ensuring that programs supported by UNICEF establish output and outcome targets and develop and implement systems for monitoring and reporting on those targets (and taking action where necessary) as early as possible in the program cycle;
- Further, working to ensure that results monitoring systems are integrated with the work of implementing partners (who are most often charged with collecting data and ensuring it is reported on at regular intervals) have the capacity and motivation to maintain and to use results monitoring systems and processes;
- Investing more effort to ensure that program theories of change developed during program design and implementation are well defined and well documented and that they support the definition of indicators for monitoring results and taking decisions; and,
- Provision of technical assistance and use of supportive supervision to sustain data collection and tracking of program results, especially by UNICEF's implementing partners.



## CHAPTER 5

# INNOVATION AND HUMAN RIGHTS-BASED APPROACHES

## CHAPTER 5

### 5.1 OBSERVATIONS ON INNOVATION

It is noteworthy that the 2016 review saw a significant increase in the number of evaluation reports that remarked on innovations in UNICEF programming when compared to the 2012 review (which identified very few). The review team members identified specific comments on technological innovations, new program models, or new implementation arrangements in 34 of the 66 evaluations reviewed.

Willingness to use innovative programming approaches or methods was the most common innovative aspect of UNICEF programming addressed in the reviewed evaluations. Examples of this type of innovation include:

- A “girl for girl” model allowing older girls to mentor their younger peers (Benin);
- The establishment of “drop in” centers for male non-school attendees in need of protection in refugee camps (Jordan);
- Innovative life skills programming to provide basic education for hard-to-reach urban working children (Bangladesh);
- Developing community maps and engaging the community in household tracking of children under five as a method of social mobilization (India);
- Use of social cash transfers to improve family capacity to care for children and improve early child development (Malawi);
- Experimenting with pay-for-performance to motivate service providers (Pakistan); and,
- Focusing on reaching men and boys specifically and targeting them for contact in “masculine spaces” such as taverns, sports events and taxi ranks (South Africa).

Evaluations also noted a number of innovations on the use of technology in UNICEF supported programs, including:

- Use of portable Point of Care (POC) machines to test for HIV antibodies (CD4 counts) among pregnant women in order to provide better case management of Anti-Retro-Viral (ART) therapies (Zimbabwe);
- Using a computer-based application to encourage at risk, school-based adolescents to take a voluntary HIV test specifically designed for their use (Namibia);
- Use of mobile phone technology for monitoring and paying teachers (Madagascar);
- Use of an on-line Child Friendly Schools (CFS) monitoring system (Turkey);
- Use of a mobile phone app as a tool for improving young people’s access to Sexual and Reproductive Health and Rights (SRHR) counselling and support services (South Africa); and,
- Use of a mobile phone based application platform for supply chain management for medicines and essential medical supplies (Malawi).

#### 5.1.1 Summary and Implications: Innovation

There are three important points which can be highlighted based on these observations regarding innovation at UNICEF:

1. Evaluations carried out by UNICEF from 2012 to 2015 are increasingly taking note of innovations and innovative aspects of UNICEF support. More than half of the evaluation reports reviewed highlighted some form of innovation in UNICEF programming. As a result, it seems clear that UNICEF is succeeding in fostering a commitment to innovation in the programs evaluated from 2012 to 2015, although there is still scope for developing a stronger culture of innovation;

2. UNICEF has been willing and able to experiment with different models of program support as a form of innovation; and,
3. The most common form of technological innovation in programs covered by UNICEF evaluations is in the incorporation of mobile phone technology, including custom made applications.

The assessment criteria used during the review did not specifically address positive and negative factors contributing to innovation in UNICEF supported programs. However, it is still possible to identify measures which could strengthen the use of innovative approaches in UNICEF programming based on the positive examples provided in the evaluations reviewed. These include:

- Arriving at a more explicit, but reasonably broad definition of innovation as a key element in UNICEF supported projects and programs;
- Developing and implementing guidance on how to more systematically document the results of innovations in programs supported by UNICEF; and,
- Ensure that innovative UNICEF programs are supported by guidelines on how to measure and document success and how to provide decision makers with evidence so that effective innovations can be taken to scale.

## 5.2 OBSERVATIONS ON HUMAN RIGHTS-BASED APPROACHES

There was a significant increase in attention paid to human rights-based approaches (HRBAP) to programming in UNICEF evaluations in the 2012 to 2015 period, when compared to 2009 to 2011. During the 2012 review, team members noted few comments on whether or not the programs under evaluation applied HRBAP principles. In contrast, for the 2016 review 39 of the 66 reviewed evaluations made a specific observation on HRBAP in UNICEF programming.

Most of the observations made are positive and point to a specific aspect of HRBAP that has been applied in the evaluated programs. The types of observations on rights included:

- Some evaluations pointed to program elements which aimed to address the knowledge, skills and attitudes of service providers (duty bearers) so that they can provide services which better address the rights of children, especially adolescents (Egypt, Colombia, Burundi, Tajikistan);
- Some focused more on the demand side, promoting the knowledge and awareness of rights holders so that they could make better use of services (South Africa, Armenia);
- Some evaluated programs tried to achieve a balance by addressing the skills, knowledge and attitudes of service providers and duty-bearers while working to empower rights holders and improve trust between them (Maldives, Palestine); and,
- A few evaluations noted that programs incorporated gender and age disaggregated data in an effort to effectively monitor and evaluate the programs contribution to strengthening the rights of children and youth (Bulgaria).

However, not all the observations on HRBAP in UNICEF supported programs were positive. Examples of negative observations in the reviewed evaluations include:

- Use of rights-based rhetoric without the application of a rights-based or gender-based analysis of the situation to identify bottlenecks to be addressed (Pakistan);
- Use of the terminology of a rights-based approach, but without any reflection of the approach in the design of interventions or any specific human rights analysis (Uzbekistan); and,
- The absence of links to other organizations working towards equity and using rights-based approaches, with the result

that a rights-based approach was not integrated into curriculum or into skills training (Bangladesh).

### 5.2.1 Summary and Implications: Rights-Based Approaches to Programming

In summary, UNICEF evaluations from 2012 to 2015 illustrate a fairly strong interest in the extent that programs are developed and implemented using a rights-based approach. When they do address this point, most evaluations are broadly positive, but a few have noted failings in the application of the principles of rights-based programming. There are identifiable steps which can improve the extent that UNICEF supported programs exhibit key characteristics of a rights-based approach. These include:

- Ensuring that UNICEF supported programs and projects balance the need to improve the skills, knowledge and attitudes of duty bearers with efforts to promote the knowledge and awareness of rights holders and to improve the level of trust between service providers and the communities they serve (balancing supply and demand side interventions); and,
- Avoiding the use of rights-based rhetoric and terminology without the application of key elements of a rights-based approach to programming including: meaningful consultations with rights holders and duty-bearers; and, a willingness to identify and address bottlenecks and barriers which limit access for the most disadvantaged.



## CHAPTER 6

# CONCLUSIONS AND RECOMMENDATIONS

## CHAPTER 6

The conclusions and recommendations presented here should not be read as indicating that UNICEF was not aware of or attempting to address the findings reported in evaluations published from 2012 to 2015. Recent and ongoing initiatives led by the Field Results Group and aimed at strengthening programming (including Results-Based Management) reported to the review team include:

- Ongoing UNICEF RBM global training efforts (2015-2017) which have reached 1,500 staff in 2016 and a planned reach of 4,000 by the end of 2017. These efforts make use of a global cadre of trainers and a suite of face-to-face and e-learning modules.
- Programme Strategy Notes – introduced in November 2015 and included in all new Country Programmes. Staff report over 85 of the Strategy Notes have been developed and each includes a theory of change, a results framework and alignment of results and strategies with expected funding. A review of the quality of the Strategy Notes prepared in 2016 is currently underway.
- The implementation of the Country Office score card with a direct link from country office performance as measured by the score card to the personal performance assessment of the UNICEF Representative.

Importantly, the terms of reference for the review and the methodology used did not include or permit any effort to assess the effectiveness of these initiatives in addressing the issues reported here.

### 6.1 CONCLUSIONS

#### Relevance

1. As was the case in 2012, the 2016 Development Effectiveness Review **demonstrates that UNICEF programming remains highly relevant**. In particular, evaluations carried out from 2012 to 2015 reported that UNICEF supported **projects and programs align well with national development priorities and programs**. While, evaluations report positive findings for relevance, they also identify **potential improvements** in the relevance of some programs through **engagement with key stakeholders during program development and through better links from needs assessment to targeted assistance**.

#### Objectives Achievement

2. **UNICEF supported programs evaluated from 2012 to 2015 have been successful in attaining their stated objectives**. The evaluations reviewed reported a strongly positive and moderately improved level of objectives achievement by UNICEF supported programs when compared to results of the 2012 review. In particular, **many evaluation reports noted UNICEF has been successful in its upstream work in support of policy and strategy development**, and in supporting effective training and capacity development programs. These are both **areas of improvement since 2012**: when the previous Development Effectiveness Review identified weak capacity development work as a negative factor which sometimes limited

objectives achievement. On the other hand, the reviewed evaluations also noted that some programs were not successful in reaching their stated objectives because of **weak program design, especially due to the absence of well-designed and clear theories of change.**

### Promoting Gender Equality

3. UNICEF's published **evaluations were much more likely to address issues of gender equality** in the period under review (2012-2015) than they did from 2009 to 2011, as reflected in the 2012 review. The higher level of attention paid to gender equality in evaluations published from 2012 to 2015 suggests that the **evaluation function at UNICEF has responded positively to the Strategic Priority Action Plan** on gender equality (2010 to 2013), as well as the results of the 2012 development effectiveness review.
4. While coverage in evaluation reports published by UNICEF has improved since the 2012 review, **the actual findings on how effectively programs address gender equality presented in those reports continue to be very disappointing.** Almost exactly the same proportion of evaluations reported positive findings on gender equality for the 2016 (48 percent) and the 2012 review (47 percent). Now, as then, **less than half the UNICEF evaluation reports which address gender equality as cross-cutting theme, report positive findings.** This apparent lack of improvement comes despite the four-year Strategic Priority Action Plan (SPAP) from 2010 to 2013.
5. When UNICEF supported programs fail to adequately address gender equality as a cross-cutting theme, by far the most common reason given in evaluation reports is the **absence of a gender perspective and/or the non-application of gender analysis during program design and delivery.** This was also the most common factor impeding effectiveness in addressing gender equality

as cited in the 2012 review. **These conclusions raise important questions about the appropriate balance between mainstreaming gender equality and the need for specific, targeted interventions aimed at addressing gender in UNICEF supported programs.**

### UNICEF's Focus on Equity

6. UNICEF evaluations published from 2012 to 2015 provide a strong indication that the **evaluation function has responded appropriately to the re-focus on equity** which began in 2010. A significant majority of the evaluations reviewed (54 of 66) directly addressed the issue of whether or not UNICEF programming effectively promotes improved equity for children and adolescents.
7. Results for this new sub-criterion on **effectiveness in improving equity were largely positive.** UNICEF supported programs were successful in addressing equity by targeting a diverse set of vulnerable or disadvantaged groups, including: poorest households by socio-economic strata; most distant, hard-to-reach rural districts; disadvantaged ethnic or tribal groups; migrant workers; children employed in urban settings and adolescent girls facing specific barriers to participation in school. It also seems clear, however, that **effectively addressing equity requires UNICEF and its partners to clearly identify the disadvantaged as a target group, to map their location and to identify the barriers which impede their access to services or assets so that these can be addressed.**
8. There remains considerable scope for improving the effectiveness of a significant portion of UNICEF supported programs in efforts to address inequality. For the most part this improvement would require UNICEF supported programs to **develop an explicit strategy to focus on benefiting marginalized individuals and communities; a strategy which is accompanied by criteria for participation and service access which provide a clear advantage to**

**the marginalized.** This strategy to address marginalized groups would be needed for upstream policy work as well as for programs directed to beneficiaries.

### Sustainability

9. As with gender equality, **the evaluation function at UNICEF has improved the level of attention paid to issues of sustainability during the 2012 to 2015 period** when compared to evaluations carried out from 2009 to 2011.
10. Despite this improved level of coverage, evaluations reviewed in 2016 continue to report **challenging findings regarding sustainability.** There has been some improvement in findings relating to strengthened community and institutional capacity as well as strengthening the policy environment since 2012. However, the basic question of whether or not the benefits of UNICEF programming will be sustained after program completion continues to generate negative findings in many evaluation reports (74 percent). In order for UNICEF supported projects and programs to improve the sustainability of results, it seems clear they will need to go beyond the noted improvements in capacity development and the enabling environment. **The evaluation reports point directly to the necessity for projects and programs to effectively address the need to plan for and ensure financial and program sustainability.**

### Efficiency

11. UNICEF programs evaluated since 2012 have been found to be **efficient in controlling costs and using resources.** They have also been successful in **implementing programs and achieving objectives on time.** However, as with the 2012 review, the evaluations point to the need for continuous attention to the cost and resource efficiency of UNICEF supported programs. They also highlight the

need to strengthen systems and processes for monitoring program costs and results in order to manage for efficiency.

### Evaluation for Improved Effectiveness

12. The review noted continuing **improvements in UNICEF's use of evaluation as a function which can contribute to increasing the effectiveness of supported programs,** including continued improvements in the quality of evaluation reports and their utility as reflected in management responses. However, there remains a **need to ensure wider coverage and a systematic approach to setting priorities for programs and projects to be evaluated so that the most relevant programs are the subject of regular, high quality, evaluations.**

### Monitoring and Results-Based Management for Improved Effectiveness

13. While UNICEF has made **substantive efforts to improve the strength of monitoring systems, and the use of results-based management (RBM) methods** since 2012, this has **not resulted in recognizable improvements** in the strength of monitoring systems from the perspective of evaluations carried out from 2012 to 2015. This may reflect the relatively recent (2014) establishment of the Field Results Group, but it also indicates that effectively strengthening monitoring and RBM systems remains a serious challenge for UNICEF.
14. Strengthening results monitoring, reporting and management at UNICEF requires the agency to engage in a **range of measures including; capacity development** for UNICEF staff and, especially, **implementing partners;** improving the clarity of **theories of change** which underpin program designs; **establishing meaningful program targets** which are **monitored using systems implemented as early as possible in the program cycle;** and, provision of **technical assistance** to sustain data collection and results tracking

by implementing partners. The mandate of the Field Results Group responds directly to many of these challenges.

## Innovation

15. Evaluations reviewed indicate that UNICEF has made **progress in fostering a commitment to innovation**, including a willingness to experiment with different models of program support and service delivery and to use emerging technologies in innovative ways. However, there are also indications that the **emphasis on innovation can be strengthened by developing a clear definition of innovative programming and focusing greater attention to programming innovations during the design of programs and services**. There is also scope for greater emphasis on documentation and taking proven innovations to scale.

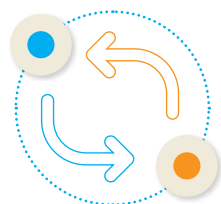
## Human Rights-based Approaches to Programming

16. UNICEF evaluations published from 2012 to 2015 provide evidence of **an increased interest in human rights-based approaches to programming**. While they indicate that many programs have made a serious attempt at a human rights-based approach they also note that some show a more rhetorical than real commitment. In particular, the evaluation reports noted **the need to address both the capacities of duty bearers to provide services and the ability of rights-holders to effectively demand those services**.

In summary, the findings of evaluations published by UNICEF in the 2012 to 2015 period, do show improvements in overall development effectiveness and in addressing key issues through evaluation. At the same time, the reviewed evaluations point to areas requiring further strengthening for improved effectiveness including improvements in programming for gender, equity and for sustainability.

## 6.2 RECOMMENDATIONS

The following recommendations are based on the findings and conclusions presented in the report. They indicate that UNICEF should:



**1** **STRENGTHEN** the capacity of UNICEF and its partners in order to **IMPROVE** program designs based on a strong theory of change.

**2**



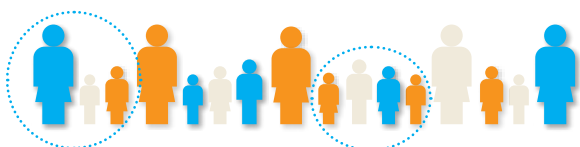
More consistently apply **GENDER ANALYSIS METHODS** and incorporate a **GENDER PERSPECTIVE** into program design and during all programming stages.

**3**

More clearly **DEFINE** and **IDENTIFY TARGET GROUPS** of disadvantaged women and children in all UNICEF supported programs and **CLARIFY** the disproportional **BENEFITS** the program will bring to the target group.

**4**

More consistently and effectively **ADDRESS PROGRAM** and **FINANCIAL SUSTAINABILITY** in the design of country programs and strategies.



**5**

Continue, and even increase efforts to **IMPROVE THE EFFICIENCY** of UNICEF supported programs.



**6**

**IMPROVE THE SELECTION OF PROGRAMS** for evaluation at UNICEF in order to ensure higher levels of program coverage and introduce more systematic decision making on priorities.



**7**

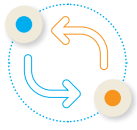


Continue and intensify efforts to **STRENGTHEN RESULTS MONITORING, REPORTING** and **MANAGEMENT SYSTEMS**.

**8**

Develop and promote a more specific definition of **INNOVATION** in UNICEF programming, one which encompasses **PROGRAMMATIC** as well as **TECHNOLOGY-BASED INNOVATION**.





# 1

## Strengthen the capacity of UNICEF and its partners in order to improve program designs based on a strong theory of change.

While UNICEF supported programs generally achieve their stated objectives, the review also identifies scope for improvement, especially in the development of meaningful theories of change which adequately specify the links from program activities and inputs to results at output and outcome levels. This should include:

- a) Capacity building of UNICEF staff and the staff of partners at country office level in the development of effective and technically sound theories of change during program design – including identifying key causal assumptions which can be tested during
- b) Capacity building and ongoing technical support to UNICEF staff and partners in the identification of meaningful program targets and appropriate indicators for tracking target achievement over time; and,
- c) Linking work on strengthening program design with efforts to strengthen results-based monitoring and management systems (see recommendation seven).

evaluations. This will require improved, concrete guidance on developing effective theories of change;



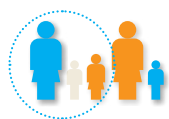
# 2

## More consistently apply gender analysis methods and incorporate a gender perspective into program design and during all programming stages.

The lack of improvement in the level of positive findings for gender equality when comparing results of the 2012 and 2016 reviews points to the need to:

- a) Ensure that program design processes and approval systems fully incorporate gender analysis and a gender perspective prior to UNICEF commitment;
- b) Strengthen the capacity of UNICEF staff and partners to make use of gender analysis tools during program design;
- c) Strengthen the collection and use of gender disaggregated data in UNICEF programmes;
- d) Engage with communities on the situation of girls and women during program design and incorporate organizations representing the interests of girls and women into stakeholder consultations;
- e) Balance the use of a gender mainstreaming approach with targeted interventions specifically aimed at addressing gender inequalities (including the inclusion of specific, clear, gender relevant objectives during the programme design phase); and,
- f) Continuing to strengthen the level of attention paid to gender equality as a cross-cutting issue in UNICEF led systems and processes for monitoring evaluation, to build on the progress made between the 2012 and 2016 reviews.

These measures are consistent with the more targeted approach adopted by UNICEF in the 2014 Gender Action Plan, especially given its strong emphasis on the need for programmes to identify and seek to address bottlenecks and barriers.



3

### More clearly define and identify target groups of disadvantaged women and children in all UNICEF supported programs and clarify the disproportional benefits the program will bring to the target group.

The results of the review indicate there is an opportunity to build on the current level of positive results by improving further the focus on equity. This should include:

- a) More consistent efforts to identify who are the most disadvantaged groups of women and children in the programming area during program design, including better mapping of their location in the communities being served;
- b) Better identification of barriers to participation and access to services for the identified groups of disadvantaged women and children so that programs specifically address them;
- c) Targeted criteria for program participation and access to services for the identified disadvantaged groups; including criteria which privilege their access and specifically anticipate a disproportionate positive result; and,
- d) Strengthened monitoring of participation, access and benefits accruing to members of the identified disadvantaged groups of women and children. This may include ensuring that selected programmes address the need to strengthen national systems for data management to better address equity.



4

### More consistently and effectively address program and financial sustainability in the design of country programs and strategies.

While there has been a notable improvement in the institutional sustainability of UNICEF supported programming, there is considerable evidence that a significant challenge to the sustainability of UNICEF supported projects and programs is the financial sustainability of the initiatives being supported. UNICEF should make every effort during program and project design and implementation to:

- a) Develop and implement mechanisms for generating income for supported projects which can be sustained after UNICEF support is completed;
- b) Where these are not appropriate, advocate effectively for host national, provincial and local government to provide necessary, line item budgets; and,
- c) Advocate for and help secure alternative sources of funding from multilateral agencies, bilateral donors or the private sector.



5

### Continue, and even increase efforts to improve the efficiency of UNICEF supported programs.

As more evaluations address issues of cost and efficiency in UNICEF supported programs, there is clearly scope to strengthen the adequacy and level of attention paid to issues of efficiency and cost effectiveness in UNICEF programming during program design. The most urgent requirement is to strengthen program designs and program monitoring systems so that they track

program costs over time and establish a credible link from unit costs to results. Otherwise, program managers cannot be expected to manage for improved efficiency. Efforts to improve efficiency could also benefit from the development of a common, core methodology for assessing program efficiency at UNICEF.

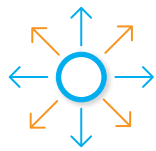


# 6

## Improve the selection of programs for evaluation at UNICEF in order to ensure higher levels of program coverage and introduce more systematic decision making on priorities.

While the recently approved Evaluation Policy recommends that all programmes should be evaluated at least once every five years, there is still scope for better linking the decision on which programs and projects should be evaluated first (given limited resources) to the requirements of both the overall agency Strategic Plan and the need to improve program effectiveness. This is especially important at the level of Country Offices, where the rationale for

selecting programs and projects for evaluation is not always clear. It may also result in greater attention being given to the links between development and humanitarian programming being supported by UNICEF. The mechanism of the Costed Evaluation Plan (CEP) may represent an important channel for strengthening the process for prioritizing evaluation decisions at country office level.



# 7

## Continue and intensify efforts to strengthen results monitoring, reporting and management systems.

This highlights the essential nature of the mandate of the Field Results Group in support of more effective results monitoring and management. In particular, there is a demonstrated need to:

- a) Develop the capacity of, in particular, UNICEF's implementing partners and their staff for designing, establishing and operating results monitoring and management systems;
- b) Improve the basic quality of project and program designs, especially as they relate to
- c) developing clear and realistic program theories of change (see recommendation one);
- d) Establish meaningful program results targets to monitor program performance against targets using systems implemented as near to the beginning of the program cycle as possible; and,
- e) Provide stronger technical assistance to sustain data collection and results monitoring and reporting by implementing partners.



# 8

## Develop and promote a more specific definition of innovation in UNICEF programming, one which encompasses programmatic as well as technology-based innovation.

Reviewed evaluations indicate that UNICEF supported programs frequently face opportunities to promote innovative methods and approaches for strengthening program design, implementation, technical support, service delivery mechanisms, and ongoing supervision and follow up. In many cases these may be as important or even more important than promoting the use of advanced technologies.

There are also opportunities to strengthen the documentation of support to innovation so that national decision makers have better information when deciding whether or not to fund UNICEF supported innovations at national scale. Improving UNICEF's modelling and evidence gathering and reporting on innovative practices could also strengthen sustainability.

Taken together, these measures should help to ensure continued improvement in the findings and conclusions reported by UNICEF's evaluation function during the life of current and succeeding agency Strategic Plans.

# ANNEXES

## ANNEX A

### DEVELOPMENT EFFECTIVENESS CRITERIA

<b>1. Relevance</b>
<ul style="list-style-type: none"> <li>1.1 Programs are suited to the needs of target group members</li> <li>1.2 Programs are aligned with national humanitarian and development goals</li> <li>1.3 Effective partnerships with government, civil society and humanitarian and development partners</li> </ul>
<b>2. Achievement of Development and Humanitarian Objectives and Expected Results</b>
<ul style="list-style-type: none"> <li>2.1 Programs and projects achieve stated objectives</li> <li>2.2 Positive benefits for target group members</li> <li>2.3 Substantial numbers of beneficiaries/contribution to national humanitarian and development goals</li> <li>2.4 Significant changes in national development policies/programs</li> </ul>
<b>3. Cross Cutting Themes: Inclusive Development and Humanitarian Which can be Sustained (Gender Equality, Environmental Sustainability and Equity)</b>
<ul style="list-style-type: none"> <li>3.1 Programs effectively address gender equality</li> <li>3.2 Changes are environmentally sustainable</li> <li>3.3 Programs effectively address the issue of equity</li> </ul>
<b>4. Sustainability</b>
<ul style="list-style-type: none"> <li>4.1 Program benefits are likely to continue</li> <li>4.2 Programs support institutional and community capacity</li> <li>4.3 Programs strengthen enabling environment for humanitarian and development</li> </ul>
<b>5. Efficiency</b>
<ul style="list-style-type: none"> <li>5.1 Program activities are cost efficient</li> <li>5.2 Programs are implemented/objectives achieved on time</li> <li>5.3 Systems for program implementation are efficient</li> </ul>
<b>6. Using Evaluation and Monitoring to Improve Humanitarian and Development Effectiveness</b>
<ul style="list-style-type: none"> <li>6.1 Systems and processes for evaluation are effective</li> <li>6.2 Systems and processes for monitoring are effective</li> <li>6.3 Results-based management systems are effective</li> <li>6.4 Evaluation results used to improve humanitarian and development effectiveness</li> </ul>

# ANNEX B

## EVALUATION SAMPLE

Title	Year	Region	Country	Outcome Area	Country Classification
2014 Afghanistan: In-depth Evaluation of Female Literacy Program	2014	ROSA	Afghanistan	Education	LIC
2015 Armenia: Evaluation of Family Support Services and Stakeholders Contribution to Related Services/Systems	2015	CEE/CIS	Armenia	Child Protection	LMIC
2015 Evaluación de la Iniciativa Maternidades Seguras y Centradas en la Familia 2010-2014	2015	LACRO	Argentina	Health	UMIC
2012 Bosnia and Herzegovina: Evaluation of the Project "Protection of Children at Risk and Children in Contact with the Justice System in Bosnia and Herzegovina"	2012	CEE/CIS	Bosnia and Herzegovina	Child Protection	UMIC
2013 Egypt: Meshwary-My Journey Project - Final Evaluation Report	2013	MENA	Egypt	Social Inclusion	LMIC
2014 Bangladesh: Final Evaluation of Basic Education for Hard to Reach Urban Working Children (BEHTRUWC) Project 2nd Phase 2004-2014	2014	ROSA	Bangladesh	Education	LIC
2013 Barbados: Determining the Results of the Koudmen Sent Lisi Pilot Program: A Social Safety Net Program in St. Lucia	2013	LACRO	Barbados	Social Inclusion	HIC
2014 Benin: Evaluation des Quatre Innovations EDUCOM au Bénin	2014	WCARO	Benin	Education	LIC
2012 Botswana: Iplegeng Evaluation	2012	ESARO	Botswana	Nutrition	UMIC
2012 Burkina Faso: Evaluation finale du projet « Approvisionnement en Eau Potable, Assainissement de base et Hygiène dans les provinces du Ganzourgou et de la Gnagna, Burkina Faso – 2007-2010»	2012	WCARO	Burkina Faso	Health	LIC
2012 Bulgaria: Evaluation of the project "Expansion of foster care model in Bulgaria"	2012	CEE/CIS	Bulgaria	Child Protection	UMIC
2013 Burundi: Evaluation of the WaSH project in the integrated rural villages of Rutana	2013	ESARO	Burundi	WASH	LIC
2015 Evaluation de da Composante Survie Et Developpement de L'enfant du Program de Cooperation Tchad-UNICEF	2015	WCARO	Chad	Health	LIC

## ANNEX B: Evaluation Sample (cont'd)

Title	Year	Region	Country	Outcome Area	Country Classification
2013 Colombia: Evaluación final del Programa Conjunto de infancia, seguridad alimentaria y nutricional	2013	LACRO	Colombia	Nutrition	UMIC
2014 Djibouti: Réhabilitation du quartier 4 de la ville de Djibouti, des systèmes d'approvisionnement en eau des villages ruraux et appui institutionnel pour la mise en oeuvre d'une politique sectorielle de l'eau	2014	MENA	Djibouti	WASH	LM
2012 Democratic Republic of Congo: Evaluation du Program Education de base 2008-2012	2012	WCARO	DR Congo	Education	LIC
2013 DRC: Evaluation externe du Program Réponse rapide aux mouvements de population (RRMP) en République Démocratique du Congo	2013	WCARO	DR Congo	Humanitarian	LIC
2013 El Salvador: Evaluación final. Ventana Temática Infancia, Seguridad Alimentaria y Nutrición	2013	LACRO	El Salvador	Nutrition	LMIC
2014 Ethiopia: An Evaluation of the Child-to-Child School Readiness Program in Ethiopia	2014	ESARO	Ethiopia	Education	LIC
2013 Ethiopia: Assessing the impact on child nutrition of the Ethiopia Community-based Nutrition Program	2013	ESARO	Ethiopia	Nutrition	LIC
2015 Fiji: Community Resilience and Coping with Climate Change and Natural Disasters in Vanuatu	2015	EAPRO	Fiji	Social Inclusion	UMIC
2013 Ghana: Evaluation of the LLIN Universal Mass Distribution "Hang-Up" Campaign in Brong Ahafo, Central and Western Regions of Ghana: Costs and Effects	2013	WCARO	Ghana	Health	LMIC
2015 Ghana: Summative Evaluation of the CI IHSS Program in Ghana	2015	WCARO	Ghana	Health	LMIC
2015 Evaluation du Projet d'Appui au Program National de Lutte contre le Choléra en Haïti	2015	LACRO	Haiti	Humanitarian	LMIC
2015 India: Evaluation of Meena Radio Program	2015	ROSA	India	Education	LMIC
2013 India: Evaluation of Social Mobilisation Network in India for Eradication of Polio	2013	ROSA	India	Health	LMIC
2012 India: Evaluation of Knowledge Community on Children in India Internship Program	2012	ROSA	India	Social Inclusion	LMIC

## ANNEX B: Evaluation Sample (cont'd)

Title	Year	Region	Country	Outcome Area	Country Classification
2014 Indonesia: Final Evaluation of the WASH in Schools program (WISE) in Indonesia	2014	EAPRO	Indonesia	WASH	LMIC
2015 Jordan: UNICEF'S Emergency Psychosocial Support Response for Syrian Children in Jordan: An Evaluation	2015	MENA	Jordan	Child Protection	UMIC
2015 Laos: Evaluation of the Water, Sanitation and Hygiene (WASH) Program of the UNICEF Country Program (2012-2014)	2015	EAPRO	Laos	WASH	LMIC
2014 Lesotho: Child Grand Impact Evaluation	2014	ESARO	Lesotho	Child Protection	LMIC
2012 Macedonia: Strengthening National Capacities to Prevent Domestic Violence	2012	CEE/CIS	Macedonia	Gender	UMIC
2015 Madagascar: Evaluation du Program d'Appui d'Urgence à l'Éducation Nationale Malgache (AUENM)	2015	ESARO	Madagascar	Education	LIC
2015 Malawi: Summative Evaluation of the CI IHSS Program in Malawi	2015	ESARO	Malawi	Health	LIC
2014 Malawi: Malawi Impact Evaluation of the National Plan of Action for Orphaned and Vulnerable Children	2014	ESARO	Malawi	Child Protection	LIC
2014 Maldives: An Evaluation of UNICEF Maldives Strategies in Addressing Issues Affecting Women and Children	2014	ROSA	Maldives	Health	UMIC
2014 Mali: Evaluation du Program WASH à l'école de l'UNICEF Mali	2014	WCARO	Mali	WASH	IC
2014 Mauritania: Evaluation indépendante de la composante survie de l'enfant du Program de coopération Mauritanie-UNICEF 2012 -2016, Période 2012-2013	2014	WCARO	Mauritania	Health	LMIC
2012 Moldova: External Evaluation of the "Child-Friendly School" Initiative (2007-2011) in the Republic of Moldova	2012	CEE/CIS	Moldova	Education	LMIC
2014 Mongolia: REDS Strategy Evaluation Mongolia	2014	EAPRO	Mongolia	Health	UMIC
2014 Montenegro: Final Evaluation of the "Child Care System Reform"	2014	CEE/CIS	Montenegro	Child Protection	UMIC
2013 Mozambique: Evaluation of the WASH Sector Strategy "Community Approaches to Total Sanitation" (CATS)	2013	ESARO	Mozambique	WASH	LIC
2014 Namibia: School Based HIV Testing and Counselling Pilot Program Evaluation	2014	ESARO	Namibia	HIV	UMIC
2015 Namibia: Impact Evaluation of the Namibia Football Association's Galz & Goals Sports for Development S4D Program	2015	ESARO	Namibia	Health	UMIC

## ANNEX B: Evaluation Sample (cont'd)

Title	Year	Region	Country	Outcome Area	Country Classification
2014 Nigeria: Impact Evaluation of Water, Sanitation, and Hygiene (WASH) within the UNICEF Country Program of Cooperation, Government of Nigeria and UNICEF, 2009-2013	2014	WCARO	Nigeria	WASH	LMIC
2013 Niger: Evaluation des Strategies du Program de Cooperation selon une Approche Basee sur l'Equite	2013	WCARO	Niger	Social Inclusion	LIC
2015 End of project Evaluation-GAVI CSO project	2015	ROSA	Pakistan	Health	LMIC
2013 Pakistan: UNICEF WASH Pakistan End of Program Evaluation for Rural Sanitation in Flood-affected Districts (RuSFAD Phase III) 2011-12	2013	ROSA	Pakistan	WASH	LMIC
2015 State of Palestine: Improving Community-Based Psychosocial Protection Services For Children And Adolescents In East Jerusalem	2015	MENA	Palestine	Child Protection	LMIC
2012 State of Palestine: Evaluation of the Child-Friendly Schools	2012	MENA	Palestine	Education	LMIC
2013 Paraguay: Evaluación del manejo e incorporación de la dimensión de igualdad de género en los programas implementados y apoyados por UNICEF/Paraguay	2013	LACRO	Paraguay	Gender	LMIC
2015 Philippines: Evaluation of the Intervention and Rehabilitation Program in Residential Facilities and Diversion Programs for Children in Conflict with the Law	2015	EAPRO	Philippines	Social Inclusion	LMIC
2013 Sierra Leone: Evaluation of integrated management of the acute malnutrition (IMAM) and infant and young child feed (IYCF) programs	2013	WCARO	Sierra Leone	Nutrition	LIC
2013 Somalia: Final Evaluation of the Unconditional Cash and Voucher Response to the 2011-2012 Crisis in Southern and Central Somalia	2013	ESARO	Somalia	Education	LIC
2012 Somalia: TOSTAN Pilot Project on "Ending FGM/C" in Northwest and Northeast Zone in Somalia	2012	ESARO	Somalia	Gender	LIC
2012 South Africa: The South African Child Support Grant Impact Assessment	2012	ESARO	South Africa	Social Inclusion	UMIC
2015 Evaluation of the DFID-funded Safer South Africa Program	2015	ESARO	South Africa	Gender	UMIC
2012 Sudan: Program Evaluation - UNICEF Assisted Water, Sanitation, and Hygiene Program In Sudan (2002-2010)	2012	MENA	Sudan	WASH	LMIC

ANNEX B: Evaluation Sample (cont'd)

Title	Year	Region	Country	Outcome Area	Country Classification
2014 Tajikistan: Juvenile Justice Alternative Project. Evaluation report	2014	CEE/CIS	Tajikistan	Child Protection	LIC
2013 United Rep. of Tanzania: Evaluation of Mobile Decision Support Applications in Tanzania	2013	ESARO	Tanzania	Health	LIC
2015 Thailand: Evaluation of the Thailand National Child and Youth Development Plan 2012-2015	2015	EAPRO	Thailand	Social Inclusion	UMC
2014 Togo: Évaluation des Interventions à Base Communautaire (Nutrition & ATPC) dans les Régions des Savanes et de la Kara	2014	WCARO	Togo	Nutrition	LIC
2015 Turkey: Evaluation of UNICEF's response to the Syrian refugee crisis in Turkey	2015	CEE/CIS	Turkey	Humanitarian	UMIC
2014 Uganda: UNICEF – Ugandan Ministry of Education and Sports BRMS Mentorship Project Evaluation	2014	ESARO	Uganda	Education	LIC
2015 Uzbekistan: Evaluation of "Improvement of Mother and Child Health Services" Phase 2	2015	CEE/CIS	Uzbekistan	Health	LMIC
2015 Evaluation of the Early Childhood Education (ECE) Interactive Radio Instruction (IRI) pilot Program in Eastern Province of Zambia	2015	ESARO	Zambia	Education	LMIC
2012 Zimbabwe: Evaluation of the Use of Point of Care PIMA CD4 Cell Count Machines for HIV Positive Women and their Families in Maternal, Newborn and Child Health (MNCH) Settings in Seven Districts in Zimbabwe	2012	ESARO	Zimbabwe	HIV	LIC

**Legend:**

- **Country Classification:** LIC – Low income country; LMIC – Low middle income country; UMIC – Upper middle income country; HIC – High income country
- **Region:** CEE/CIS – Central and Eastern Europe and Commonwealth of Independent States; ESARO – Eastern and Southern Africa; EAPRO – Eastern Asia and the Pacific; LACRO – Latin American and the Caribbean; ROSA – South Asia; WCARO – Western and Central Africa

# ANNEX C

## METHODOLOGY

### The Evaluation Population and Sample

The review team identified 391 evaluation reports uploaded to the UNICEF data base of published evaluations in the period from the beginning of 2012 to the end of 2015. This population was the starting point for a sampling process which aimed to produce a purposive sample of approximately 70 evaluations which could be reviewed by the team and their findings classified by the criteria presented in section 3.2. The size of the sample was selected based on experience in completed DERs, especially relating to the fact that some evaluation criteria and sub-criteria are covered less frequently than others. With a review encompassing approximately 70 evaluation reports, experience indicates that almost all of the criteria and sub-criteria will be covered by a sufficient number of evaluations to allow the team to make a credible statement on the distribution of findings across the four levels used by the review (from highly unsatisfactory to highly satisfactory).

The process used to draw the purposive sample began with a series of steps aimed at defining the **sample frame**. These included:

1. Identification of the universe of 391 evaluations uploaded to the data base from 2012 to 2015;
2. Elimination of evaluation reports receiving a rating of Unsatisfactory from UNICEF's Global Evaluation Report Oversight (GEROS) system;
3. Elimination of global thematic evaluations and regional evaluations in order to avoid double counting and duplication of results with country level evaluations;
4. Elimination of case studies, mid-term and process evaluations, interim studies and reviews which are unlikely to address many of the OECD DAC evaluation criteria; and,
5. Elimination of joint evaluations and UNDAF evaluations, where UNICEF programming may not be addressed directly.

The resulting sample frame of country level evaluations of UNICEF's development and humanitarian effectiveness consisted of 263 evaluation reports. Given that some evaluation reports would be excluded during the process of quality review carried out by the review team, an initial sample of 80 evaluation reports was drawn from the sample frame (30.4 percent).

The process of drawing the actual sample began with the random selection of 80 evaluation reports stratified by country (to allow at least a reasonable distribution of evaluations among UNICEF program countries). This initial stratified random sample was then adjusted by eliminating specific evaluation reports that contributed to the over and under representation of a specific outcome area, region, year of publication or GEROS rating (when comparing the sample to both the population of evaluations and the distribution of UNICEF program expenditures). In all, a total of ten evaluations were eliminated from the stratified random sample and replaced by ten others.

The key characteristics of the original sample in comparison to the evaluation report sample frame and the distribution of UNICEF Program expenditures are illustrated in a series of tables in Annex C.

Of the 80 evaluation reports in the original sample, only 73 were subjected to the internal quality review by the DER team. The seven reports not reviewed were de-selected for the following reasons:

- One on closer examination was a 2011 report;
- One was a baseline study rather than an effectiveness evaluation;
- One was written and published in Portuguese;
- Four were partially reviewed but were excluded because they were narrowly focused on criteria which could not be organized around the six core criteria of the Development Effectiveness Review.

Of the remaining 73 evaluation reports, a further seven (9.6 percent) were excluded from the analysis due to quality scores which were lower than the minimum score determined in advance (25 of 40 available points). Interestingly, the rate at which evaluation reports were excluded for quality reasons was higher for the 2015 DER than for the one undertaken in 2012. During the previous DER, four of 66 evaluation reports (6.0 percent) were excluded for failing the quality review

The quality review criteria, the individual weighting scores for each criterion and the detailed results of the quality review process, including the distribution of total scores, are all presented in Annex D.

The final sample reviewed for this DER included 66 evaluation reports (25 percent of those in the sample frame of 263). It has the following characteristics illustrated in detail in Annex C:

- The distribution of evaluation reports in the sample is very similar to the distribution of reports by **year of publication** in the overall population, with just over 30 percent published in 2015;
- The distribution of sampled reports by UNICEF Program region is very close to the regional distribution of reports in the sample frame; the CEE/CIS region is slightly under-represented accounting for 13.6 percent of reports in the sample, and 17.5 percent of those in the frame;

- The sample tracks very closely to the distribution of UNICEF evaluation reports by level of country development with a slight over-representation of lower-middle-income countries at 39.4 percent of the sample and 34.2 percent of the population;
- In order to better represent the budgetary balance of funding across defined output areas of UNICEF programming, the sample slightly over-represents the health sector with 21.2 percent of the sample versus 12.5 percent of the population. Similarly, it over-represents the humanitarian output area at 6.1 percent of the sample, this partly reflects the very low number of evaluation reports in the population with a direct focus on predominantly humanitarian operations (just 2.7 percent). It is important to note that the evaluation universe, and the sample, does include evaluations classified under other output areas, for example water and sanitation or health programming, which take place in a humanitarian or emergency context and so have potential implications for humanitarian programming. However, the actual review process of evaluations included in the sample, did not provide enough evidence of finding on humanitarian operations to support conclusions on humanitarian programming.

In summary, the resulting sample of 66 evaluation reports (Annex C) represents a modified (and thus purposive), stratified random sample of approximately 25 percent of the evaluation reports included in the sample frame. It is meant to be illustrative of the results of UNICEF's efforts to assess development and humanitarian effectiveness during the period 2012 to 2015. Broadly speaking, the goal of the sample is to present a faithful reflection of UNICEF's own findings and conclusions regarding effectiveness of its programs as evaluated in the field.

## Evaluation Criteria

As with all DERs carried out since 2010, the methodology and analytical approach of this Development Effectiveness Review focuses on essential characteristics of multilateral organizations as derived from the OECD DAC Evaluation Criteria. As a result, the review methodology focused on six main criteria and 20 sub-criteria considered to be the core elements of effective developmental and humanitarian programming.

Nineteen of the 20 effectiveness sub-criteria described in Table 1, are identical to the criteria used in the 2013, DER of UNICEF which covered

evaluations published from 2008 to 2011. One has been added to improve the relevance of the review to UNICEF's current programming.

1. Sub-criteria 3.3 requires the review to identify and classify reported evaluation findings which relate to the extent to which UNICEF supported programs effectively address issues of equity and, thereby respond to a key strategic direction for the organization. The guidelines for analysts when classifying evaluation findings relating to equity can be found in Annex I under sub-criteria 3.3.

**TABLE 3**      **Criteria for Assessing Developmental and Humanitarian Effectiveness**

<b>5 Relevance</b>
<ul style="list-style-type: none"> <li>1.1 Programs are suited to the needs of target group members</li> <li>1.2 Programs are aligned with national humanitarian and development goals</li> <li>1.3 Effective partnerships with government, civil society and humanitarian and development partners</li> </ul>
<b>6 Achievement of Development and Humanitarian Objectives and Expected Results</b>
<ul style="list-style-type: none"> <li>2.1 Programs and projects achieve stated objectives</li> <li>2.2 Positive benefits for target group members</li> <li>2.3 Substantial numbers of beneficiaries/contribution to national humanitarian and development goals</li> <li>2.4 Significant changes in national development policies/programs</li> </ul>
<b>7 Cross Cutting Themes: Inclusive Development and Humanitarian Which can be Sustained (Gender Equality, Environmental Sustainability and Equity)</b>
<ul style="list-style-type: none"> <li>3.1 Programs effectively address gender equality</li> <li>3.2 Changes are environmentally sustainable</li> <li>3.3 Programs effectively address the issue of equity</li> </ul>
<b>8 Sustainability</b>
<ul style="list-style-type: none"> <li>4.1 Program benefits are likely to continue</li> <li>4.2 Programs support institutional and community capacity</li> <li>4.3 Programs strengthen enabling environment for humanitarian and development</li> </ul>
<b>9 Efficiency</b>
<ul style="list-style-type: none"> <li>5.1 Program activities are cost efficient</li> <li>5.2 Programs are implemented/objectives achieved on time</li> <li>5.3 Systems for program implementation are efficient</li> </ul>
<b>10 Using Evaluation and Monitoring to Improve Humanitarian and Development Effectiveness</b>
<ul style="list-style-type: none"> <li>6.1 Systems and processes for evaluation are effective</li> <li>6.2 Systems and processes for monitoring are effective</li> <li>6.3 Results-based management systems are effective</li> <li>6.4 Evaluation results used to improve humanitarian and development effectiveness.</li> </ul>

It is important to note that this addition has served to extend the analysis and enrich the information gathered from the reviewed evaluation reports. It does not diminish the comparability of results across the two completed DERs, because all 19 sub-criteria from the 2013 study were used in the current round.

## Review Process and Data Analysis

Each evaluation report in the sample was reviewed by a member of a small review team of two reviewers and two senior members (including the team leader). The first task of each reviewer was to assess the quality of the evaluation report under examination using the Quality Review Grid (Annex D). The reviewers classified the evaluation findings in each evaluation report as they apply to the six evaluation criteria and 20 sub-criteria presented in Table 1. Each finding was classified using a web-based evaluation review reporting tool under one of five different headings:

1. Not Addressed
2. Highly Unsatisfactory
3. Unsatisfactory
4. Satisfactory and
5. Highly Satisfactory

The guidelines for classifying each evaluation finding have been developed for each evaluation sub-criteria and each level of finding (from highly unsatisfactory to highly satisfactory). The detailed guidelines are presented in Annex I.

In order to strengthen intra-reviewer consistency and reliability, the review process began with a two-day calibration and training session including the analysts and senior team members who reviewed the same evaluation report and discussed and calibrated their findings. The team also held regular meetings to discuss criteria and compare findings during the review process. The team leader and deputy-team leader also reviewed completed evaluation review reports during the process to

correct any misalignment and convene discussions of the team where there were any questions regarding the interpretation of evidence.

During each review, the analysts not only classified evaluation findings, they entered into the web-based review grid for each evaluation, the key evidence used to classify evaluation findings. This allowed the senior team members to assess the accuracy with which all findings have been classified and coded during the review.

Prior to statistical profiling of the results of the review, senior team members reviewed all ratings to ensure there was sufficient evidence to support the rating and that it was consistent with the guidelines. The draft final report combines the quantitative evidence from the 66 evaluation reports reviewed, with qualitative evidence gathered from the review of a sample of thematic evaluations and from interviews with UNICEF staff. This approach is entirely consistent with the data gathering and analysis process used in the 2012 review and in all other completed DERs since 2010.

## Limitations

As a meta-synthesis of previously completed evaluation reports, this development effectiveness review has been subject to a set of limitations. The most important of these include:

- **Sampling Bias:** The evaluation sample, while it began with a stratified random sample, was adjusted to purposely improve representation of evaluations in the health sector and evaluations from countries in sub-Saharan Africa. This resulted in a purposive sample that was intended to be illustrative of the results of UNICEF evaluations published from 2012 to 2015. It is not a statistically valid random sample of UNICEF evaluations which can be used to calculate confidence intervals around a set of point estimates.
- **Bias in the Evaluation Universe:** While there is a very significant body of evaluation reporting by UNICEF in the period, the sample frame of evaluations under-represents some

geographic areas and program outcomes as identified in the strategic plan. A meta-synthesis of evaluation results is inherently limited in coverage to that achieved by the population of evaluation reports it samples from. Any biases in the evaluation universe can be expected to be reflected, to some degree and despite efforts to balance representation, in the evaluation sample.

- **Limits to Stratification:** The evaluation sample has been balanced to be largely reflective of the distribution of evaluation reports produced by UNICEF in relation to geographic area, year of publication and outcome area of the strategic plan. However, the resulting strata of the sample are too small to allow for meaningful reporting of results in each sub-component (the different outcome areas for example). However, it has been possible to compare some results across broader aggregations of stratification (for example, between least developed countries (LDCs) and other categories of countries).
- **Coverage of Humanitarian Programming:** Limits to stratified analysis in the sample have particular importance when considering the humanitarian dimension of UNICEF's work. Because evaluations of humanitarian programming accounted for only 2.7 percent of all UNICEF published evaluations during the review period, a very small number could be included in the review if the sample was to reflect the overall population. While humanitarian focused evaluations were somewhat oversampled (they made up 6.1 percent of the sample), they still were too limited in number to allow any generalization of results to UNICEF's overall experience in humanitarian programming. It is important to note that the evaluation universe, and the sample, both include evaluations classified

under other output areas, for example water and sanitation or health programming, which take place in a humanitarian or emergency contexts and so have potential implications for humanitarian programming. However, the actual review process of evaluations included in the sample, did not provide enough evidence of finding on humanitarian operations to support conclusions on humanitarian programming.

- **The Retrospective Nature of Meta-Synthesis:** Evaluation reports are, by definition, retrospective in that they provide ex-ante assessments of the effectiveness of programs. Many of the programs evaluated from 2012 to 2015 originated some years before they were evaluated. Further, publication of the findings may only have occurred some time after the evaluation field work was completed. This means that the review needed to take care not to hold older programs accountable for new strategies or to generalize historical results to current practice.

While these limitations are real, none strongly undermine the utility of the results of the development effectiveness review process. By taking care in selecting the evaluation sample and in classifying and analyzing evaluation results, the review team has been able to replicate the 2012 process and provide a meaningful overview of the existing evaluative evidence of organizational effectiveness (and how it has changed in the intervening period). The process provides UNICEF with one of the very few approaches available for aggregating field level, ground proven findings on effectiveness. Further, it allows UNICEF to take advantage of investments in evaluation which have been already made, by organizing and analyzing the results of those investments in a systematic way.

# ANNEX D

## COMPARISON OF POPULATION AND SAMPLE

### Distribution by Year

Year	Evaluation Population (n= 263)		Final Sample Analyzed (n= 66)	
	Number	%	Number	%
2015	75	28.5%	20	30.3%
2014	58	22.1%	18	27.3%
2013	66	25.1%	15	22.7%
2012	64	24.3%	13	19.7%

### Distribution by Region

Year	Evaluation Population (n= 263)		Final Sample Analyzed (n= 66)	
	Number	%	Number	%
CEE/CIS	46	17.5%	9	13.6%
EAPRO	28	7.7%	6	9.1%
ESARO	56	29.2%	19	28.8%
LACRO	26	9.2%	6	9.1%
MENA	23	9.2%	6	9.1%
ROSA	28	12.3%	8	12.1%
WCARO	56	18.5%	12	18.2%

### Distribution by Country Classification

Year	Evaluation Population (n= 263)		Final Sample Analyzed (n= 66)	
	Number	%	Number	%
LIC	104	39.5%	23	34.8%
LMIC	90	34.2%	26	39.4%
UMIC	60	22.8%	17	25.8%
HI	4	1.5%	-	-
N/A	5	1.9%	-	-

## ANNEX D: Comparison of Population and Sample (cont'd)

**Distribution by Outcome Area**

Year	Evaluation Population (n= 263)		Final Sample Analyzed (n= 66)	
	Number	%	Number	%
Health	33	12.5%	14	21.2%
HIV&AIDS	15	5.7%	2	3.0%
WASH	34	12.9%	9	13.6%
Nutrition	15	5.7%	5	7.6%
Education	59	22.4%	11	16.7%
Child Protection	41	15.6%	9	13.6%
Social Inclusion	24	9.1%	8	12.1%
Humanitarian <sup>33</sup>	7	2.7%	4	6.1%
Gender	8	3.0%	4	6.1%
Not Classified	27	10.3%	-	-

**Comparison of the Sample with Program Expenditure by Region**

Year	Program Expenses (2014) - \$4.13 billion (% by region)	Final Sample Analyzed – 66 (% by region)
Sub-Saharan Africa	56.7%	47.0%
Asia-Pacific	18.3%	21.2%
Latin America and the Caribbean	4.0%	9.1%
Middle East and Northern Africa	14.3%	9.1%
CEE/CIS	2.8%	13.6%
Inter-regional	3.9%	-

**Legend:****Regions**

CEE/CIS – Central and Eastern Europe and Commonwealth of Independent States;

ESARO – Eastern and Southern Africa;

EAPRO – Eastern Asia and the Pacific;

LACRO – Latin American and the Caribbean;

ROSA – South Asia;

WCARO – Western and Central Africa

**Country Classification**

LIC – Low income country;

LMIC – Low middle income country;

UMIC – Upper middle income country;

HIC – High income country

<sup>33</sup> It is important to note that evaluation reports in the other output areas will also, in some cases, cover humanitarian programming.

# ANNEX E

## EVALUATION QUALITY REVIEW GRID AND RESULTS

	Points for criteria scored	Maximum Points	Score
<b>A</b>	<b>Purpose of the evaluation:</b> <ul style="list-style-type: none"> <li>- why the evaluation was done (1)</li> <li>- what triggered the evaluation (including timing in the project/ program cycle) (1)</li> <li>- how evaluation is to be used (1)</li> </ul>	3	
<b>B</b>	<b>Evaluation objectives</b> <ul style="list-style-type: none"> <li>- evaluation objectives are clearly stated (1)</li> <li>- objectives logically flow from purpose (1)</li> </ul>	2	
<b>C</b>	<b>Organization of the evaluation</b> <ul style="list-style-type: none"> <li>- logical structure to the organization of the evaluation (1)</li> <li>- evaluation is well written (1)</li> <li>- clear distinction between evidence, findings, conclusions, and recommendations (1)</li> </ul>	3	
<b>D</b>	<b>Subject evaluated is clearly described</b> Evaluation describes: <ul style="list-style-type: none"> <li>- the activity/program being evaluated (1)</li> <li>- the program's expected achievements (for humanitarian: identification of needs of target population and agency specific role in addressing needs) (1)</li> <li>- how the program addresses the development problem (1)</li> <li>- the implementation modalities used (1)</li> </ul>	4	
<b>E</b>	<b>Scope of the evaluation</b> Evaluation defines the boundaries of the evaluation in terms of: <ul style="list-style-type: none"> <li>- time period covered (1)</li> <li>- implementation phase under review (1)</li> <li>- geographic area (1)</li> <li>- dimensions of stakeholder involvement being examined (1)</li> </ul>	4	
<b>F</b>	<b>Evaluation criteria</b> Evaluation criteria include: <ul style="list-style-type: none"> <li>- the achievement of development objectives and expected results (including impacts) (1)</li> <li>- cross-cutting issues: inclusive development which is gender sensitive (1)</li> <li>- cross-cutting issues: inclusive development which addresses equity issues of children and adolescents (1)</li> <li>- the sustainability of benefits and positive results achieved (1)</li> <li>- the relevance of UNICEF activities and supported projects and programs (1)</li> <li>- the efficiency of UNICEF operations in support of projects and programs (1)</li> </ul>	6	
<b>G</b>	<b>Multiple lines of evidence</b> <ul style="list-style-type: none"> <li>- one point (1) for each line of evidence used (case studies, surveys, site visits, interviews, etc.), up to a maximum of four points (4)</li> </ul>	4	

## ANNEX E - Evaluation Quality Review Grid and Results (cont'd)

	Points for criteria scored	Maximum Points	Score
<b>H</b>	<p><b>Evaluation design</b> Elements of a good evaluation design include:</p> <ul style="list-style-type: none"> <li>- a theory of how objectives and results were to be achieved (1)</li> <li>- specification of the level of results achieved (output, outcome, impact) (1)</li> <li>- baseline data (quantitative or qualitative) on conditions prior to program implementation and/or comparison of conditions after program delivery to those before (1)</li> <li>- a qualitative or quantitative comparison of conditions among program participants and a control group (not applicable to humanitarian evaluation, assign 1 point) (1)</li> </ul>	4	
<b>I</b>	<p><b>Evaluation findings and conclusions are relevant and evidence based</b> Evaluation report includes:</p> <ul style="list-style-type: none"> <li>- evaluation findings relevant to the assessment criteria (1)</li> <li>- findings that are supported by the chosen methodologies (1)</li> <li>- a clear logical link between the evidence and the finding (1)</li> <li>- conclusions which are linked to the evaluation findings as reported (1)</li> </ul>	4	
<b>J</b>	<p><b>Evaluation limitations</b></p> <ul style="list-style-type: none"> <li>- statement of the limitations of the methodology (1)</li> <li>- impact of limitations on evaluation (1)</li> <li>- remedies of limitations (1)</li> </ul>	3	
<b>K</b>	<p><b>Evaluation Recommendations</b></p> <ul style="list-style-type: none"> <li>- evaluation contains recommendations that flow from findings and conclusions (1)</li> <li>- recommendations are directed to one or more organization (1)</li> <li>- recommendations are aimed at improving development effectiveness (1)</li> </ul>	3	
	<b>Total (required to have a minimum of 25 points)</b>	<b>40</b>	
	<b>Total for Criteria G, H and I (required to have minimum of 9 points)</b>	<b>12</b>	

The quality review was completed for 73 of the 80 reports originally selected for review. Three of them were not reviewed because they were part of the earlier exclusion criteria – baseline, 2011 report and report in Portuguese. Another four were not reviewed / partially reviewed and dropped because the reports were written in a way that was not suitable for the scoring criteria of the DER. The result of the quality screening is presented in the following table.

Three of the eleven quality assessment criteria (G, H and I) are directly linked to the quality of evidence provided since they focus on: a) the

use of multiple lines of evidence (Quality Assessment Criterion G); b) the strength of the evaluation design including an explicit program theory, identified results expectations, use of baselines, and comparison tests (Quality Assessment Criterion H; and, c) the relevance and evidence based nature of findings (Quality Assessment Criterion I).

These three criteria are interlinked and all speak to the quality of the evaluation report and the reliability of findings and substantiating evidence.

## Results of Quality Review Scoring

Overall evaluation quality scores			Development effectiveness key criteria (G, H & I)		
Max points = 40 (Min required = 25)	No. of evaluations	% of evaluations	Max points = 12 (Min required = 9)	No. of evaluations	% of evaluations
36 - 40	18	24.7%	12	11 (1)	15.1%
31 - 35	36	49.3%	11	21	28.8%
25 - 30	13 (1)	17.8%	10	22 (1)	30.1%
21 - 24	(4)	5.5%	9	15 (1)	20.5%
16 - 20	(1)	1.4%	8	(1)	1.4%
11 - 15	(1)	1.4%	7	(2)	2.7%
6 - 10	-	-	1-6	(1)	1.4%
<b>Total</b>	<b>73 (7)</b>	<b>100.0%</b>	<b>Total</b>	<b>66</b>	<b>100.0%</b>

It is true that these three criteria account for 12 of the 40 points available for the overall quality score but for that very reason, the DER approach requires that each report receive a minimum of nine of the available 12 points for design related criteria G, H and I (a 75% requirement). It is for this reason that some evaluations were excluded from the analysis. This approach is consistent with the method used to date in all DERs

Out of the total 73 reports reviewed for quality, seven were rejected based on quality scores. The figures in brackets are the number of reports that failed quality screening. From the table above it could be seen that there was one

report that had a total score of 25 or more but it failed because it did not meet the minimum score of nine in G, H, and I Criteria. At the same time, three reports that got 10 or more on the development effectiveness criteria G, H, I criteria did not get the minimum total.

Five of the seven evaluations that failed were from WCARO region or from earlier years of the DER – 2012 (3) and 2013 (2)

Reports (66 of them) that scored 25 or more and a minimum score of nine of on development effectiveness criteria G, H and I were considered passed on quality. These were used to complete the review and used for analysis.

# ANNEX F

## GLOBAL EVALUATIONS AND CORPORATE DOCUMENTS REVIEWED

### 1. Global Evaluations Sample

- Global Evaluation of UNICEF's Peace Building, Education and Advocacy (PBEA) – 2015
- Protecting Children from Violence – A Comprehensive Evaluation of UNICEF's Strategies and Program Performance – 2015
- Evaluation of UNICEF's Upstream Work in Basic Education and Gender Equality (2003-2012) - 2014
- Evaluation of UNICEF's Cluster Lead Agency Role in Humanitarian Action – 2013
- Evaluation of UNICEF's Emergency Preparedness Systems – 2013
- Global Evaluation of Life Skills Education Programs – 2012
- Global Evaluation of the Application of the Human Rights-Based Approach (HRBA) to UNICEF – 2012
- UNICEF CEE/CIS Multi-Country Evaluations – Synthesis Report - 2015

### 2. Corporate Documents

- UNICEF Strategy and Policy Documents
- Revised Evaluation Policy of UNICEF (2013)
- UNICEF Evaluation Policy (2008)
- UNICEF's Revised Evaluation Policy: Evolving Strategy Issues – Presentation by Evaluation Office, November 13, 2014
- UNICEF Strategic Plan 2014-2017
- Final results framework of the UNICEF Strategic Plan 2014-2017
- Synthesis Report on the Evaluation Policy and Function – Final Report, February 2013

### UNICEF Reports

- Annual reports of Executive Director on performance and results (for 2012, 2013 and 2014)
- Annual reports on the evaluation function (for 2012, 2013 and 2014)
- Draft annual report on the evaluation function for 2015
- Management perspective on the annual reports on the valuation function (2012, 2013, and 2014)
- UNICEF Annual Reports (2012, 2013 and 2014)
- UNICEF Strategic Plan 2014-2017 Evaluability Assessment – Final Report, July 2015
- Quadrennial Comprehensive Policy Review of operational activities for development of the United Nations system – Report of Secretary General, 2012
- Annual Results Report – Water Sanitation and Hygiene, 2015
- Annual Results Report – Humanitarian Action, 2014 and 2015
- The State of the World's Children 2015 - Reimagine the future: Innovation for every Child

### GEROS Documents

- External Assessment of UNICEF's Global Evaluation Reports Oversight System (GEROS) – Report, January 2016
- Global Evaluation Report Oversight System (GEROS) – Global Meta-Evaluation Reports (2013, 2014, 2015)

### **Other UNICEF Documents**

- MOPAN report (2012)
- System-Wide Action Plan Analysis – 2013
- System-Wide Action Plan Analysis - 2014
- Review of UNICEF’s Development Effectiveness (2009-2011) – Final Report, April 30, 2013
- Formative Evaluation of UNICEF’s Monitoring Results for Equity System (MoRES) – Evaluation Report, August 2014
- Taxonomy for Defining and Classifying UNICEF Research, Evaluation and Studies, UNICEF Office of Research, September 2014
- KPI Definitions and Grading Criteria – FRG internal document
- Performance Scorecard Methodology and Indicator Framework – FRG internal document
- Demonstrating Results – FRG presentation December 10, 2015
- UNICEF Mozambique Costed Evaluation Plan 2017-2020

## ANNEX G

### LIST OF STAKEHOLDERS INTERVIEWED

#### Corporate Headquarters, New York

1. Colin Kirk, Director of Evaluation, Evaluation Office
2. Samuel Bickel, Senior Adviser, Research and Evaluation, Evaluation Office
3. Mathew Varghese, Senior Evaluation Specialist, Evaluation Office
4. Ada Ocampo, Senior Evaluation Specialist, Systemic Strengthening, Evaluation Office
5. Laurence Reichel, Evaluation Specialist, Evaluation Office
6. Stephen Adkisson, Associate Director, Field Results Group
7. Timothy Takona, Senior Adviser, Field Results Group
8. George Laryea-Adjei, Deputy Director, Division of Data, Research and Policy
9. Stefan Peterson, Associate Director and Chief of Health Section, Program Division,
10. Mathieu Brossard, Senior Advisor, Education
11. Luwei Pearson, Senior Advisor, Health
12. Josephine Bourne, Associate Director and Chief of Education, Program Division
13. Cornelius Williams, Associate Director Child Protection
14. Jan Werner Schultink, Associate Director Nutrition
15. Evariste Kouassi Komlan, Senior Adviser, Water Sanitation and Hygiene
16. Yasmin Ali Haque, Deputy Director, Office of Emergency Programmes
17. Dominique Hyde, Deputy Director, Public Partnership Division
18. Vidya Ganesh, Deputy Director, Program Division (scheduled)
19. Manzia Zaman-Cespedes, Representative-Tanzania- MoRES
20. Anju Malhotra, Principal Adviser, Gender, Rights and Development
21. Alexandra Yuster, Associate Director, Social Inclusion

#### Regional Offices

1. Roumiana Gantcheva – Chief of M&E, MENARO
2. Hermani Mirella, M&E Specialist, LACRO
3. Bastiaan Van't Hoff, Chief of Program and Planning, LACRO
4. Inoussa Kabore, Chief of M&E, WCARO
5. Edward Addai, Chief of M&E, ESARO
6. Lori Bell, Regional M&E Adviser, CEE/CIS
7. Riccardo Polastro, Regional Evaluation Adviser. EAPRO
8. Urs Nagel. Regional Evaluation Adviser, ROSA

# ANNEX H

## COMPARISON OF RESULTS BY COUNTRY TYPE

1.1 UNICEF supported programs and projects are suited to the needs and/or priorities of the target group

	HUS	US	S	HS	Total	Valid #	S & HS	Total
LIC	-	35.3%	47.1%	17.6%	100%	17	64.7%	23
LMIC	-	30.4%	47.8%	21.7%	100%	23	69.6%	26
UMIC	-	25.0%	56.3%	18.8%	100%	16	75.0%	17

1.2 UNICEF supported development projects and programs align with national development goals:

	HUS	US	S	HS	Total	Valid #	S & HS	Total
LIC		5.6%	16.7%	77.8%	100%	18	94.4%	23
LMIC	4.2%		54.2%	41.7%	100%	24	95.8%	26
UMIC			42.9%	57.1%	100%	14	100.0%	17

1.3 UNICEF has developed an effective partnership with governments, bilateral and multilateral development and humanitarian organizations and NGOs for planning, coordination and implementation of support to development...

	HUS	US	S	HS	Total	Valid #	S & HS	Total
LIC	8.3%	25.0%	58.3%	8.3%	100%	12	66.7%	23
LMIC	0.0%	33.3%	52.4%	14.3%	100%	21	66.7%	26
UMIC	0.0%	23.1%	46.2%	30.8%	100%	13	76.9%	17

2.1 UNICEF supported programs and projects achieve their stated development and/or humanitarian objectives and attain expected results.

	HUS	US	S	HS	Total	Valid #	S & HS	Total
LIC	4.8%	9.5%	61.9%	23.8%	100%	21	85.7%	23
LMIC	0.0%	20.0%	64.0%	16.0%	100%	25	80.0%	26
UMIC	5.9%	41.2%	47.1%	5.9%	100%	17	52.9%	17

2.2 UNICEF supported programs and projects have resulted in positive benefits for target group members.

	HUS	US	S	HS	Total	Valid #	S & HS	Total
LIC	0.0%	0.0%	78.3%	21.7%	100%	23	100.0%	23
LMIC	-	8.0%	88.0%	4.0%	100%	25	92.0%	26
UMIC	-	6.3%	81.3%	12.5%	100%	16	93.8%	17

2.3 UNICEF programs and projects made differences for a substantial number of beneficiaries and where appropriate contributed to national development goals.

	HUS	US	S	HS	Total	Valid #	S & HS	Total
LIC	-	20.0%	60.0%	20.0%	100%	20	80.0%	23
LMIC	-	17.4%	78.3%	4.3%	100%	23	82.6%	26
UMIC	-	53.3%	33.3%	13.3%	100%	15	46.7%	17

## ANNEX H - Comparison of Results by Country Type (cont'd)

## 2.4 UNICEF 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...

	HUS	US	S	HS	Total	Valid #	S & HS	
LIC	-	22.2%	61.1%	16.7%	100%	18	77.8%	23
LMIC	-	17.4%	56.5%	26.1%	100%	23	82.6%	26
UMIC	-	14.3%	71.4%	14.3%	100%	14	85.7%	17

## 3.1 Extent UNICEF supported activities effectively address the cross-cutting issue of gender equality.

	HUS	US	S	HS	Total	Valid #	S & HS	
LIC	11.1%	33.3%	55.6%	0.0%	100%	18	55.6%	23
LMIC	4.3%	43.5%	52.2%	0.0%	100%	23	52.2%	26
UMIC	-	64.7%	35.3%	0.0%	100%	17	35.3%	17

## 3.3 Extent UNICEF programming contributes to equity

	HUS	US	S	HS	Valid	NA	Total	
LIC	-	53.8%	38.5%	7.7%	100%	13	46.2%	23
LMIC	0.0%	31.8%	68.2%	0.0%	100%	22	68.2%	26
UMIC	14.3%	28.6%	57.1%	0.0%	100%	14	57.1%	17

## 4.1 Benefits continuing or likely to continue after project or program completion or there are effective measures to link the humanitarian operations, to rehabilitation, reconstructions...

	HUS	US	S	HS	Total	Valid #	S & HS	
LIC	22.2%	55.6%	22.2%	0.0%	100%	18	22.2%	23
LMIC	8.7%	52.2%	39.1%	0.0%	100%	23	39.1%	26
UMIC	0.0%	53.3%	46.7%	0.0%	100%	15	46.7%	17

## 4.2 Extent UNICEF supported projects and programs are reported as sustainable in terms of institutional and/or community capacity.

	HUS	US	S	HS	Total	Valid #	S & HS	
LIC	5.3%	31.6%	52.6%	10.5%	100%	19	63.2%	23
LMIC	0.0%	43.5%	52.2%	4.3%	100%	23	56.5%	26
UMIC	0.0%	21.4%	71.4%	7.1%	100%	14	78.6%	17

## 4.3 Extent UNICEF development programming contributes to strengthening the enabling environment for development.

	HUS	US	S	HS	Total	Valid #	S & HS	
LIC	0.0%	53.8%	38.5%	7.7%	100%	13	46.2%	23
LMIC	0.0%	25.0%	60.0%	15.0%	100%	20	75.0%	26
UMIC	6.7%	33.3%	60.0%	-	100%	15	60.0%	17

## 5.1 Program activities are evaluated as cost/resource efficient:

	HUS	US	S	HS	Total	Valid #	S & HS	
LIC	0.0%	25.0%	50.0%	25.0%	100%	16	75.0%	23
LMIC	0.0%	40.9%	50.0%	9.1%	100%	22	59.1%	26
UMIC	7.1%	35.7%	35.7%	21.4%	100%	14	57.1%	17

## ANNEX H - Comparison of Results by Country Type (cont'd)

### 5.2 Evaluation indicates implementation and objectives achieved on time (given the context, in the case of humanitarian programming)

	HUS	US	S	HS	Total	Valid #	S & HS	
LIC	-	28.6%	64.3%	7.1%	100%	14	71.4%	23
LMIC	0.0%	46.7%	46.7%	6.7%	100%	15	53.3%	26
UMIC	0.0%	40.0%	60.0%	0.0%	100%	10	60.0%	17

### 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...)

	HUS	US	S	HS	Total	Valid #	S & HS	
LIC	10.0%	60.0%	20.0%	10.0%	100%	10	30.0%	23
LMIC	9.5%	52.4%	38.1%	0.0%	100%	21	38.1%	26
UMIC	0.0%	54.5%	27.3%	18.2%	100%	11	45.5%	17

### 6.1 Systems and process for evaluation are effective.

	HUS	US	S	HS	Total	Valid #	S & HS	
LIC	0.0%	20.0%	75.0%	5.0%	100%	20	80.0%	23
LMIC	0.0%	31.8%	54.5%	13.6%	100%	22	68.2%	26
UMIC	0.0%	42.9%	50.0%	7.1%	100%	14	57.1%	17

### 6.2 Systems and processes for monitoring and reporting on program results are effective

	HUS	US	S	HS	Total	Valid #	S & HS	
LIC	9.5%	47.6%	33.3%	9.5%	100%	21	42.9%	23
LMIC	13.0%	47.8%	26.1%	13.0%	100%	23	39.1%	26
UMIC	12.5%	56.3%	25.0%	6.3%	100%	16	31.3%	17

### 6.3 Results-Based Management (RBM) systems are effective

	HUS	US	S	HS	Total	Valid #	S & HS	
LIC	19.0%	42.9%	38.1%		100%	21	38.1%	23
LMIC	12.5%	45.8%	37.5%	4.2%	100%	24	41.7%	26
UMIC	18.8%	56.3%	0.0%	25.0%	100%	16	25.0%	17

### 6.4 MO makes use of evaluation to improve development/humanitarian effectiveness

	HUS	US	S	HS	Total	Valid #	S & HS	
LIC	13.0%	0.0%	17.4%	69.6%	100%	23	87.0%	23
LMIC	23.1%	3.8%	15.4%	57.7%	100%	26	73.1%	26
UMIC	17.6%	0.0%	35.3%	47.1%	100%	17	82.4%	17

### 6.5 Management response results in significant action taken

	HUS	US	S	HS	Total	Valid #	S & HS	
LIC	13.0%	0.0%	17.4%	69.6%	100%	23	87.0%	23
LMIC	23.1%	15.4%	57.7%	3.8%	100%	26	61.5%	26
UMIC	17.6%	5.9%	64.7%	11.8%	100%	17	76.5%	17

# ANNEX I

## COMPARISON OF RESULTS OF UNICEF DER 2016 AND 2012

1.1 UNICEF supported programs and projects are suited to the needs and/or priorities of the target group

	HUS	US	S	HS	Total	Valid #	S & HS	Total
DER 2012	5%	5%	58%	32%	100%	60	90%	62
DER 2016	-	30%	50%	20%	100%	56	70%	66

1.2 UNICEF supported development projects and programs align with national development goals:

	HUS	US	S	HS	Total	Valid #	S & HS	Total
DER 2012	-	2%	56%	43%	100%	61	98%	62
DER 2016	2%	2%	39%	57%	100%	56	96%	66

1.3 UNICEF has developed an effective partnership with governments, bilateral and multilateral development and humanitarian organizations and NGOs for planning, coordination and implementation of support to development...

	HUS	US	S	HS	Total	Valid #	S & HS	Total
DER 2012	5%	7%	57%	30%	100%	56	87%	
DER 2016	2%	28%	53%	17%	100%	46	70%	66

2.1 UNICEF supported programs and projects achieve their stated development and/or humanitarian objectives and attain expected results.

	HUS	US	S	HS	Total	Valid #	S & HS	Total
DER 2012	7%	16%	61%	16%	100%	61	77%	62
DER 2016	3%	22%	59%	16%	100%	63	75%	66

2.2 UNICEF supported programs and projects have resulted in positive benefits for target group members.

	HUS	US	S	HS	Total	Valid #	S & HS	Total
DER 2012	3%	11%	65%	21%	100%	62	86%	62
DER 2016	-	5%	82%	13%	100%	64	95%	66

2.3 UNICEF programs and projects made differences for a substantial number of beneficiaries and where appropriate contributed to national development goals.

	HUS	US	S	HS	Total	Valid #	S & HS	Total
DER 2012	5%	23%	41%	31%	100%	61	72%	62
DER 2016	-	28%	60%	12%	100%	58	72%	66

2.4 UNICEF 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...

	HUS	US	S	HS	Total	Valid #	S & HS	Total
DER 2012	12%	12%	40%	37%	100%	43	77%	62
DER 2016	-	18%	62%	20%	100%	55	82%	66

## ANNEX I - Comparison of Results of UNICEF DER 2016 and 2012 (cont'd)

### 3.1 Extent UNICEF supported activities effectively address the cross-cutting issue of gender equality.

	HUS	US	S	HS	Total	Valid #	S & HS	Total
DER 2012	14%	38%	40%	7%	100%	42	47%	62
DER 2016	5%	47%	48%	-	100%	58	48%	66

### 3.3 Extent UNICEF programming contributes to equity

	HUS	US	S	HS	Total	Valid #	S & HS	Total
DER 2012	-	-	-	-	-	-	-	NA
DER 2016	4%	29%	65%	2%	100%	54	67%	66

### 4.1 Benefits continuing or likely to continue after project or program completion or there are effective measures to link the humanitarian operations, to rehabilitation, reconstructions...

	HUS	US	S	HS	Total	Valid #	S & HS	Total
DER 2012	12%	37%	44%	7%	100%	59	51%	62
DER 2016	11%	53%	36%	-	100%	56	36%	66

### 4.2 Extent UNICEF supported projects and programs are reported as sustainable in terms of institutional and/or community capacity.

	HUS	US	S	HS	Total	Valid #	S & HS	Total
DER 2012	17%	36%	36%	12%	100%	59	48%	62
DER 2016	2%	34%	57%	7%	100%	56	64%	66

### 4.3 Extent UNICEF development programming contributes to strengthening the enabling environment for development.

	HUS	US	S	HS	Total	Valid #	S & HS	Total
DER 2012	10%	20%	63%	8%	100%	40	71%	
DER 2016	2%	35%	55%	8%	100%	48	73%	66

### 5.1 Program activities are evaluated as cost/resource efficient:

	HUS	US	S	HS	Total	Valid #	S & HS	Total
DER 2012	-	38%	54%	8%	100%	48	62%	62
DER 2016	2%	35%	46%	17%	100%	52	63%	66

### 5.2 Evaluation indicates implementation and objectives achieved on time (given the context, in the case of humanitarian programming)

	HUS	US	S	HS	Total	Valid #	S & HS	Total
DER 2012	18%	31%	46%	5%	100%	39	51%	62
DER 2016	-	38%	57%	5%	100%	39	62%	66

### 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...

	HUS	US	S	HS	Total	Valid #	S & HS	Total
DER 2012	17%	46%	34%	3%	100%	35	37%	62
DER 2016	7%	55%	31%	7%	100%	42	38%	66

## ANNEX I - Comparison of Results of UNICEF DER 2016 and 2012 (cont'd)

## 6.1 Systems and process for evaluation are effective.

	HUS	US	S	HS	Total	Valid #	S & HS	Total
DER 2012	5%	31%	45%	19%	100%	42	64%	62
DER 2016	-	30%	61%	9%	100%	56	70%	66

## 6.2 Systems and processes for monitoring and reporting on program results are effective

	HUS	US	S	HS	Total	Valid #	S & HS	Total
DER 2012	12%	49%	27%	12%	100%	59	39%	62
DER 2016	12%	50%	28%	10%	100%	60	38%	66

## 6.3 Results-Based Management (RBM) systems are effective

	HUS	US	S	HS	Total	Valid #	S & HS	Total
DER 2012	na	na	na	na	Na	24	24	62
DER 2016	16%	48%	28%	8%	100%	61	36%	66

## 6.4 MO makes use of evaluation to improve development/humanitarian effectiveness

	HUS	US	S	HS	Total	Valid #	S & HS	Total
DER 2012	38%	5%	21%	36%	100%	61	57%	62
DER 2016	18%	2%	21%	59%	100%	66	80%	66

## 6.5 Management response results in significant action taken

	HUS	US	S	HS	Total	Valid #	S & HS	Total
DER 2012	-	-	-	-	-	-	-	NA
DER 2016	18%	11%	63%	8%	100%	66	71%	66

# ANNEX J

## ANALYSIS GUIDE

CRITERIA	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
<b>Common Development and Humanitarian Evaluation Assessment Criteria</b>				
<b>1. Relevance</b>				
1.1 UNICEF supported programs and projects are suited to the needs and/or priorities of the target group	Evaluation finds that substantial elements of program or project activities and outputs were unsuited to the needs and priorities of the target group.	Evaluation finds that no systematic analysis of target group needs and priorities took place during the design phase of developmental or humanitarian 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.	Evaluation finds that the UNICEF 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.	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).
1.2 UNICEF supported humanitarian and development projects and programs align with national humanitarian and development goals:	The evaluation reports that significant elements of UNICEF supported humanitarian and development program and project activity run counter to national humanitarian and development priorities with a resulting loss of effectiveness.	The evaluation reports a <b>significant portion</b> (1/4 or more) of the UNICEF 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.	<b>Most</b> UNICEF 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 and sector plans and priorities. Wherever multilateral organization supported programs and projects are reported in the evaluation as not directly supportive of national plans and priorities they do not run counter to those priorities.	<b>All</b> UNICEF 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.

## ANNEX J - Analysis Guide (cont'd)

CRITERIA	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
1.3 UNICEF 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 response and recovery efforts.	Evaluation finds that UNICEF 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 UNICEF 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 UNICEF 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 UNICEF has consistently achieved a high level of partnership during the evaluation period.
<b>2. Achievement of Humanitarian and Development Objectives and Expected Results</b>				
2.1 UNICEF supported programs and projects achieve their stated development and/or humanitarian objectives and attain expected results.	Less than half of stated output and outcome objectives have been achieved including one or more very important output and/or outcome level objectives.	Half or less than half of stated output and outcome level objectives are achieved.	UNICEF supported programs and projects either achieve at least a majority of stated output and outcome objectives (more than 50% if stated) or that the most important of stated output and outcome objectives are achieved.	UNICEF supported programs and projects achieve all or almost all significant development and/or humanitarian objectives at the output and outcome level.
2.2 UNICEF supported programs and projects have resulted in positive benefits for target group members.	Problems in the design or delivery of UNICEF supported activities mean that expected positive benefits for target group members have not occurred or are unlikely to occur.	UNICEF 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.	UNICEF supported projects and programs have resulted in positive changes experienced by target group members (at the individual, household or community level). These benefits may include the avoidance or reduction of negative effects of a sudden onset or protracted emergency.	UNICEF 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 the avoidance or reduction of negative effects of a sudden onset or protracted emergency.

ANNEX J - Analysis Guide (cont'd)

CRITERIA	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
<p>2.3 UNICEF programs and projects made differences for a substantial number of beneficiaries and where appropriate contributed to national development goals.</p>	<p>UNICEF supported projects and programs have not contributed to positive changes in the lives of beneficiaries as measured quantitatively or qualitatively.</p>	<p>UNICEF 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).</p>	<p>UNICEF 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, humanitarian operations and may include the avoidance of negative effects of emergencies.</p>	<p>UNICEF 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 response and recovery objectives agreed to with the national government and/or national and international development and relief organizations.</p>
<p>2.4 UNICEF activities contributed to significant changes in national development policies and programs (including for disaster preparedness, emergency response) (policy impacts) and/or to needed system reforms.</p>	<p>National policies and programs in a given sector or area of development (including disaster preparedness, emergency response and recovery) were deficient and required strengthening but UNICEF activities have not addressed these deficiencies.</p>	<p>UNICEF 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 recovery. (Policy changes in humanitarian situations may include allowing access to the effected populations).</p>	<p>UNICEF 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 recovery.</p>	<p>UNICEF 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 recovery. Further, the supported policies and program implementation modalities have resulted in improved positive impacts for target group members.</p>
<p><b>3. Cross Cutting Themes: Inclusive Humanitarian Assistance and Development Which can be Sustained</b></p>				
<p>3.1 Extent UNICEF supported activities effectively address the cross-cutting issue of gender equality.</p>	<p>UNICEF supported activities are unlikely to contribute to gender equality or may in fact lead to increases in gender inequalities.</p>	<p>UNICEF supported activities either lack gender equality 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).</p>	<p>UNICEF supported programs and projects achieve a majority (more than 50%) of their stated gender equality objectives.</p>	<p>UNICEF supported programs and projects achieve all or nearly all of their stated gender equality objectives.</p>

## ANNEX J - Analysis Guide (cont'd)

CRITERIA	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
<p>3.2 Extent changes are environmentally sustainable.</p>	<p>UNICEF supported programs and projects do not include planned activities or project design criteria intended to promote environmental sustainability. In addition changes resulting from UNICEF supported programs and projects are not environmentally sustainable.</p>	<p>UNICEF 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.</p> <p>OR</p> <p>UNICEF supported programs and projects include planned activities or project design criteria intended to promote sustainability but these have not been successful.</p>	<p>UNICEF 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.</p>	<p>UNICEF 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.</p>
<p>3.3 The extent to which UNICEF supported programs effectively address issues of equity for children and adolescents.</p>	<p>The evaluation reports that the program under review failed to identify or take note of barriers to equity based on any one of gender, poverty, ethnicity, geographic isolation, disability, or urban/rural divide. The program need not have addressed all of these but should have taken note of one or more. Where the program identified barriers to equity, no specific effort was included in the design to address them.</p>	<p>The evaluation reports that the program identified one or more barriers to equity based on gender, poverty, ethnicity, etc. but the program features designed to help marginalized or under-served groups overcome them have been shown to be inadequate or have not resulted in any improvement in access.</p>	<p>The evaluation reports that the program identified one or more significant barriers to equity and incorporated specific measures to help target group members overcome them. Further, these measures have resulted to some extent in improvements in equity of access to and use of services and/or a more equitable distribution of positive results.</p>	<p>The evaluation reports that the program identified one or more significant barriers to equity and incorporated measures to help target group members overcome them. Further, the measures have resulted in significant improvements in equity of access to and use of services and a more equitable distribution of results. They have also strengthened the ability of target group members to demand equitable treatment and hold service providers accountable.</p>

ANNEX J - Analysis Guide (cont'd)

CRITERIA	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
<b>4. Sustainability</b>				
4.1 Benefits continuing or likely to continue after project or program completion or there are effective measures to link the humanitarian operations, to recovery, 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 project completion. For humanitarian operations, the evaluation finds no strategic or operational measures to link relief, to recovery, reconstruction and, eventually, to development.	There is a low probability that the program/project will result in continued benefits for the target group after completion. For humanitarian operations, efforts to link the relief phase to recovery, reconstruction and, eventually, to development are 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)	Likely that the program or project will result in continued benefits for the target group after completion. For humanitarian operations, the strategic and operational measures to link relief to recovery, reconstruction and, eventually, development are credible.	Highly likely that the program or project will result in continued benefits for the target group after completion. For humanitarian operations, the strategic and operational measures to link relief to recovery, reconstruction and, eventually, development are credible. Further, they are likely to succeed in securing continuing benefits for target group members.
4.2 Extent UNICEF supported projects and programs are reported as sustainable in terms of institutional and/or community capacity.	The design of UNICEF 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 recovery and/or development.	UNICEF 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 recovery and/or development.	UNICEF programs and projects have contributed to strengthening institutional and/or community capacity but with limited success.	Either UNICEF 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.

## ANNEX J - Analysis Guide (cont'd)

CRITERIA	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
4.3 Extent UNICEF development programming contributes to strengthening the enabling environment for development.	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, service delivery and quality; and necessary improvements to supporting structures such as capital and labour markets). Further, the UNICEF activities and support provided to programs and projects failed to address the identified weakness successfully, further limiting program results.	UNICEF development activities and/or UNICEF supported projects and programs have not made a notable contribution to changes in the enabling environment for development.	UNICEF development activities and/or UNICEF 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, service delivery and quality; and necessary improvements to supporting structures such as capital and labour markets.	UNICEF development activities and/or UNICEF 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, service delivery and quality; and necessary improvements to supporting structures such as capital and labour markets. Further, these improvements in the enabling environment are leading to improved development outcomes.
<b>5. Efficiency</b>				
5.1 Program activities are evaluated as cost/resource efficient:	Credible information indicating that UNICEF supported programs and projects (development, emergency preparedness, relief and recovery) are not cost/resource efficient.	UNICEF supported programs and projects under evaluation (development, emergency preparedness, relief and recovery) 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.	Level of program outputs achieved (development, emergency preparedness, relief and recovery) 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.	UNICEF supported (development, emergency preparedness, relief and recovery) programs and projects are designed to include activities and inputs that produce outputs in the most cost/resource efficient manner available at the time.

ANNEX J - Analysis Guide (cont'd)

CRITERIA	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
5.2 Evaluation indicates implementation and objectives achieved on time (given the context, in the case of humanitarian programming)	Less than half of stated output and outcome level objectives of UNICEF supported programs and projects are achieved on time, there is no credible plan or legitimate explanation found by the evaluation which would suggest significant improvement in on-time objectives achievement in the future.	Less than half of stated output and outcome level objectives of UNICEF supported programs and projects are achieved on time but the program or project design has been adjusted to take account of difficulties encountered and can be expected to improve the pace of objectives achievement in the future. In the case of humanitarian programming, there was a legitimate explanation for the delays.	More than half of stated output and outcome level objectives of UNICEF supported programs and projects are achieved on time and that this level is appropriate to the context faced by the program during implementation, particularly for humanitarian programming.	Nearly all stated output and outcome level objectives of UNICEF 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 outputs/outcomes is provided.
5.3 Evaluation indicates that UNICEF 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.)	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.	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.	Agency systems and procedures for project implementation are reasonably efficient and have not resulted in significant delays or increased costs.	Efficiency of agency systems and procedures for project implementation represent an important organizational strength in the implementation of the program under evaluation.
<b>6. Using Evaluation and Monitoring to Improve humanitarian and development Effectiveness</b>				
6.1 Systems and process for evaluation are effective.	Evaluation practices in use for programs and projects of this type (development, emergency preparedness, relief and recovery) are seriously deficient.	No indication that programs and projects of this type (development, emergency preparedness, relief and recovery) are subject to systematic and regular evaluations.	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 recovery programming.	Program being evaluated (along with similar programs and projects) is subject to systematic regular evaluations or describes significant elements of such practice.

## ANNEX J - Analysis Guide (cont'd)

CRITERIA	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
6.2 Systems and processes for monitoring and reporting on program results are effective	Absence of monitoring and reporting systems for the development and humanitarian programming. This would include the absence of adequate monitoring of outputs during the implementation of humanitarian programming.	While monitoring and reporting systems for the development and humanitarian programming exist, they either do not report on a regular basis or they are inadequate in frequency, coverage or reliability.	Monitoring and reporting systems for development and humanitarian programming as appropriate are well-established and report regularly.	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 are reportedly used in the management of the program.
6.3 Results-Based Management (RBM) systems are effective	No evidence of the existence of an RBM system for the program and no system is being developed.	While an RBM system is in place, or being developed, it is unreliable and does not produce regular reports on program performance.	RBM system is in place and produces regular reports on program performance.	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.
6.4 UNICEF makes use of evaluation to improve development/humanitarian effectiveness	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.	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.	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.	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.

# ANNEX K

## TERMS OF REFERENCE

### UNICEF Evaluation Office

#### Development Effectiveness Review of UNICEF 2012-2014

##### Terms of Reference

### 1. BACKGROUND

UNICEF is mandated by the United Nations General Assembly 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. The organization is also committed to continually improve its performance and part of this commitment includes ensuring the effectiveness of its development programmes.

In recent years there has been increased attention to development effectiveness not just at UNICEF but at all UN agencies, funds and programmes and among donors. The Quadrennial Comprehensive Policy Review of operational activities for development of the United Nations system (including UNICEF) includes a focus on development effectiveness and UNICEF's own Strategic Plan 2014 – 2017 also emphasizes the need to continue the agency's work in this area.

To learn more about UNICEF's development effectiveness Evaluation Office is commissioning a review that will synthesize findings from past evaluations. The lessons identified through exercise will inform ongoing policy, strategy and programme development and help the agency best achieve results for children. This document outlines the purpose, scope, methodology and use of the review and describes the team UNICEF is seeking to conduct it.

### 2. OBJECTIVE AND SCOPE

#### General objective

The present Request for Proposals. (REP) concerns the conduct of a Development Effectiveness Review (DER) of UNICEF. The review is being commissioned by the UNICEF Evaluation Office.

The purpose of the Development Effectiveness Review is to generate a body of credible information on a common set of criteria that provides an insight into the development effectiveness of UNICEF. The review builds on information, in particular evaluation reports, which is already available from the organization. It applies methods that are modest in time and costs and is to be conducted with a limited burden on UNICEF.

This DER will be the second in a series. The first was completed in 2013 by the Government of the Netherlands on behalf of the OECD countries. It covered the period 2009-2011. It is the objective of this DER to make comparisons with the first to see if **UNICEF** is maintaining its performance in areas rated satisfactory or better in the first exercise, as well as improving upon those where it was rated less well.

#### Scope

The review entails conducting a systematic synthesis of information from available evaluations published in 2012-2015 as they relate to the criteria of development effectiveness. The analysis includes also a meta-analysis of the quality of the evaluations.

#### Intended Utilization

This report will be used in 3 ways:

- A) By UNICEF managers, to gain an independent external view of UNICEF's development effectiveness. UNICEF commissions such reviews from time to time and participates in

those conducted by external bodies such as the Multilateral Organizations Performance Assessment Network (MOPAN). Such reviews inform decisions on issues such as adjustments to results-based management systems, and determining which sectors or areas of activity need reinforcement to become more effective.

- B) By members of the UNICEF Executive Board and UNICEF partners, especially those requiring evidence of how well UNICEF is performing. This review is one of the commitments made to and accepted by the Executive Board in the global plan of thematic evaluations discussed at the 2/2014 Board session.
- C) By stakeholders in the UNICEF Evaluation Function. The primary data source for the DER will be completed evaluations conducted by UNICEF. The review will therefore provide a clear window into the performance of the evaluation function. This DER will be an important part of the information base for external reviews or a peer review exercise.

### 3. TOPICS TO BE COVERED IN THE REVIEW

The review will supply evidence to judge how well UNICEF meets these 6 over-arching effectiveness criteria:

- A) Are UNICEF's programs relevant?
- B) Do the programs supported by UNICEF in the development and humanitarian spheres achieve their expected results?
- C) Does UNICEF properly address the cross-cutting themes of gender equality and environmental sustainability?
- D) Are the UNICEF supported programs creating conditions for long-term sustainability?
- E) Are the programs implemented in an efficient manner?

- F) Does UNICEF use evaluation and monitoring to improve humanitarian and development effectiveness?

The specific sub-criteria (questions) are visible in the final report issued in 2013. For ease of reference, use the link given in section 4 next, and consult Annexes 1 and 4 for full details on the questions and the criteria.

## 4. METHODOLOGY AND APPROACH

### Work Plan/Technical Proposal

The key portion of the technical proposal in response to this RFP will be a work plan that outlines an approach and methodology to accomplish the objectives. The size of the work plan should be no longer than 20 pages.

Bidders need not construct the methodology from scratch. The initial exercise and this one are to be based on this reference document: *Assessing the Development Effectiveness of Multilateral Organizations; Guidance on the Methodological Approach* (OECD/DAC Network on Development Evaluation, Paris, June 2012).

When developing the approach and work plan, bidders are expected to emphasize how they will address these issues within their overall design:

- A) Ensuring that the evaluations to be synthesized are selected (sampled) in order to adequately cover UNICEF's policies and specific themes therein. It is suggested to cover evaluations published in the period 2012 – 2015 and placed on UNICEF's Global Evaluation Database (website). This database contains various types of evaluation including: country program evaluations; program evaluations; regional and global thematic evaluations; project evaluations; process evaluations. It contains evaluations conducted/commissioned by UNICEF's Evaluation Office and evaluations conducted/commissioned by the regional or country offices or other entities of UNICEF.

- B) Determining the quality of the evaluations to be reviewed. The quality of the evaluation reports is annually assessed in UNICEF through the Global Evaluation Report Oversight System (GEROS). The results of these assessments are contained in annual GEROS reports (which include those covering evaluations completed in 2012, 2013, and 2014). This information can be taken into account in the Development Effectiveness Review. At the same time, the guidance document outlines a process for determining whether the subject organization's (UNICEF's) own quality assurance processes (i.e. GEROS) are sufficiently robust that the sampled materials need to be rated again within the DER or not. This 2016 DER exercise will follow the same methodology as the 2012-2013 exercise in this regard.
- C) Maintaining a continuity with the approach and results of the first development effectiveness exercise. In the simplest instance, this involves using the same methodology to the degree possible and arriving at conclusions that can be directly compared with those of the earlier exercise. At a more advanced level, it means understanding the findings of the first exercise, and checking to see if there have been improvements along the lines implied by the findings. (The word 'implied' is used specifically as the report did not make formal recommendations. In the absence of formal recommendations, UNICEF did not provide a management response).
- D) Defining how the team expects to manage its interaction with UNICEF (Evaluation Office and possibly other entities) during the review in a way that places the minimum necessary burden on UNICEF.
- E) Organizing the work among the team members. The approach/methodology (work plan) of the Development Effectiveness Review of UNICEF should provide information on: 1) the division of tasks between

the respective team members: team leader, senior analyst, junior analysts / IT specialist; 2) the Quality Assurance (QA) arrangements to be applied during the review *inter ea* focusing on how possible inter-researcher bias can be reduced/avoided. Please note that, as detailed in the RFP documentation that accompanies this TOR, the expected level of effort for each named person must be presented in the technical proposal, though without the attendant costs which are only given in the financial proposal.

For items A, B and C in particular, demonstrating the capacity to replicate or to extend the methodology of the previous review is essential.

### Reference Materials

In responding to this tender, bidders will be helped, and at some points are required, to relate to and build upon a similar exercise that was conducted in 2012-2013 by the Government of the Netherlands acting on behalf of the OECD countries. The full final report is available and constitutes an important resource when developing the proposal. In addition, an important guidance document on the methodology is also available.

Other background material that might be of interest to prospective bidders may be found on the general UNICEF website, [www.unicef.org](http://www.unicef.org). This includes information on the programming sectors of the present strategic plan, on the evaluation function in the organization, etc.

## 5. MANAGEMENT AND GOVERNANCE ARRANGEMENTS

### General Governance

This exercise is to be supervised by the Evaluation Office of UNICEF-NY Headquarters. The evaluation team will report to a senior member of the Evaluation Office

A Reference Group of key stakeholders will be formed to act as a support body to the Evaluation office and the evaluation team. The main functions of the reference group will be to react to the deliverables, to guide the evaluation team in understanding UNICEF, and to develop confidence in the overall exercise that will lead to a willingness to use the report findings once delivered.

### Risk Management

Many of the normal sources of risk are not expected to occur in this exercise. There will be no visits to field office settings. There will be no direct collection of data from program beneficiaries. Hence, travel/security and ethical risks are considered to be negligible.

The main risks are thought to be those of slow/weak responsiveness by UNICEF to evaluator requests for information, and potential losses to key personnel in the evaluation team to illness of other factors. The Evaluation Office has the accountability to arrange timely engagement and responsiveness from UNICEF. The organization winning the bid has the responsibility to nominate qualified substitutes—subject to UNICEF approval—should a team member need to be replaced for any reason.

Other risks that might be revealed during the inception phase will be jointly discussed at that time.

## 6. TEAM COMPOSITION

The Development Effectiveness Review of UNICEF will be carried out by a small team of independent assessors. The team should contain at least 4 named persons: a team leader/senior analyst, a senior analyst and two analysts. IT-support is required to handle the data generated during the review. Bidders may propose a larger but not a smaller team.

The team will work under the management oversight of the organization submitting the bid. Individuals may not apply as individuals, but must be within a team in a bid submitted by an incorporated body.

The team should adhere to the following specifications:

The **team leader** should have:

- a university degree (MA) in one of social sciences or in evaluation, statistics, or research methods.
- at least 10 years of experience with international development assistance;
- at least one of the provided references (which can include supervisors on prior assignments) includes the role as team leader for multidisciplinary evaluation teams;
- at least one of the provided references includes experience in evaluation of development assistance at policy and programme level;
- evaluation experience in complex evaluations and meta-analysis;
- evaluation expertise and experience of multi-lateral development assistance;
- excellent writing and communication skills;
- fluency in English;
- be available throughout the period during which the review will be conducted.

The **senior analyst** should have:

- a university degree (MA) in one of social sciences or in evaluation, statistics, or research methods.
- at least 10 years of experience with international development assistance;
- at least one reference that can describe experience as a member of multidisciplinary evaluation teams;

- at least one reference that can describe experience in evaluation of development assistance at policy and programme level;
- evaluation experience in complex evaluations and meta-analysis;
- evaluation expertise and experience of multi-lateral development assistance;
- excellent writing and communication skills (English and at least working knowledge of French and/or other UN languages);
- should be able to replace the team leader if required.

The **analysts** should have:

- a university degree in social sciences, humanities or economics;
- knowledge of international development assistance;
- knowledge of evaluation design and methodology;
- experience with literature/document search and analysis;
- good writing skills;
- fluency in English and/or French and/or other UN languages.

The team should include **an IT specialist** to handle the data generated during the review process. One of the analysts may double as IT specialist. In such a case he/she should have experience in data handling including the necessary statistical skills.

The team as a whole will be advantaged to the degree it can show these characteristics as a group:

- a) Thematic evaluation expertise across multiple of the sectors contained in UNICEF's strategic plan, including Health, Education, HIV-AIDS, Nutrition, WASH, Child Protection, and Social Policy. Likewise, cross-cutting expertise in Communication for

Development, in Gender, and in rights-based/equitybased programmes are advantages.

- b) Programming and/or evaluation experience in the diverse contexts in which UNICEF operates, including but not limited to the range extending from least developed nations to middle-income countries; in fragile states and humanitarian contexts of different types of crises; and in distinct language/culture zones.

## 7. DELIVERABLES

### Major Deliverables

- A) Inception Report: to contain a fully elaborated methodology, including, inter alia:
  - a) The sampling strategy for reading completed evaluations in the time period under review and for any other data gathering efforts
  - b) The fully described tools to be employed
  - c) An evaluation matrix relating the methods to be used to the questions to be answered
  - d) A final report outline, with any intended differences from the first DER highlighted
  - e) A confirmed time table for completing the work; and,
  - f) Confirmation of team roles and quality assurance processes
- B) A draft final report that will be the basis of discussion with UNICEF on the findings and recommendations.
- C) Final report, including an executive summary of 8-10 pages and a main report of up to 70 pages, plus annexes. The report is expected to include material setting out comparisons and contrasts with the previous DER and to address issues which may not have been covered in the previous report.

## **PHOTO CREDITS**

Chapter 1: © UNICEF/UNI94241/Bito

Chapter 2: © UNICEF/UNI182963/Noorani

Chapter 3: © UNICEF/UNI175127/Vishwanathan

Chapter 4: © UNICEF/UNI189328/Gilbertson V

Chapter 5: © UNICEF/UNI201826/Schermbrucke

Chapter 6: © UNICEF/UN05993/Anmar



