



United Nations Children Fund (UNICEF) Development Effectiveness Review, 2016- 2019

Evaluation Office

September, 2020

Table of Contents

Acronym	i
Executive Summary	ii
1. Introduction	1
1.1 Background and History of Effectiveness Reviews at UNICEF	1
1.2 Purpose.....	1
2. Applying Review Results to the SP Evaluation.....	3
2.1 Highlights of the Strategic Plan	3
3. Methodology and Assessment Criteria	4
3.1 Methodology: A Brief Overview	4
3.2 Evaluation Coverage	5
3.3 Limitations	7
4. Evaluation Findings.....	8
4.1 Relevance.....	8
4.2 Achievement of Objectives	15
4.3 Cross-Cutting Themes	22
4.4 Sustainability	29
4.5 Efficiency.....	33
4.6 Using Evaluation and Monitoring to Improve Effectiveness.....	38
5. Innovation.....	44
6. Evaluation of the Strategic Plan and Implications.....	47
7. Conclusions.....	55
Annexes.....	60
Annex A: Criteria for Assessing Developmental and Humanitarian Effectiveness.....	60
Annex B: Evaluation Sample (62 evaluations).....	61
Annex C: Methodology	66
Annex D: Comparison of Population and Sample	71
Annex E: Evaluation Quality Review Grid and Results.....	74

Table of Contents

Annex F: Global Evaluations and Corporate Documents Reviewed.....	78
Annex G: Key UNICEF Staff Consulted.....	79
Annex H: Comparison of Results by Country Type	81
Annex I: Guide for Review Team to Classify Evaluation Findings	85
Annex J: Changes Over Time Weighted Average Scores 2012, 2016 and 2020	93

Tables

Table 1: Level of Coverage for Each Evaluation Criteria	6
Table 2: Examples of Meeting Target Group Needs.....	10
Table 3: Examples of Needs Assessment Strategies	12
Table 4: Examples of Alignment	13
Table 5: Examples of Objectives Achievement by UNICEF Goal Area	18
Table 6: Supporting Changes to National Policies and Programs.....	21
Table 7: Contributing to Gender Equity in UNICEF Goal Areas	24
Table 8: Contributions to Addressing Inequities.....	25
Table 9: Reaching Children with Disabilities.....	26
Table 10: Contributing to Equity for Adolescents.....	27
Table 11: Distribution of Evaluation Quality Scores 2020, 2016 and 2012	42
Table 12: Strategic Plan Change Strategies and Applicable Effectiveness Criteria	49
Table 13: DER Findings and Baseline for Change Strategies and Enablers	51

Figures

Figure 1: Results for Relevance Criteria	9
Figure 2: Results for Objectives Achievement Criteria.....	17
Figure 3: Results for Cross Cutting Themes.....	23
Figure 4: Results for Sustainability.....	31
Figure 5: Results for Efficiency.....	35
Figure 6: Results for Evaluation and Monitoring.....	39
Figure 7: Goal Areas, Change Strategies, Organizational Enablers of the UNICEF Strategic Plan	49

Acronym

CS	Change Strategy
CSO	Civil Society Organization
DAC	Development Assistance Committee
DAPM	Division of Data, Analytics, Planning and Monitoring
DER	Development Effectiveness Review
ECD	Early Childhood Development
EN	Organizational Enabler
GAP	Gender Action Plan
GBV	Gender-Based Violence
HIV	Human Immunodeficiency Virus
HPV	Human Papillomavirus
INGO	International Non-Governmental Organization
LSE	Life Skills Education
M&E	Monitoring and Evaluation
MNCAH	Maternal Newborn Child and Adolescent Health
MTSP	Medium Term Strategic Plan
NGO	Non-Governmental Organization
OECD	The Organisation for Economic Cooperation and Development
DER	Development Effectiveness Review
RBM	Results-based Management
RO	Regional Office
SDG	Sustainable Development Goal
SGBV	Sexual Gender-Based Violence
UN	The United Nations
UNDAF	United Nations Development Assistance Framework
UNEG	United Nations Evaluation Group
UNICEF	United Nations Children Fund
UNPAF	United Nations Partnership Agreement Framework
WASH	Water, Sanitation and Hygiene

Executive Summary

Background and Purpose

This is the third consecutive development effectiveness review (DER) undertaken at UNICEF to synthesize findings from its own published evaluation reports. The first was commissioned by the Ministry of Foreign Affairs of the Netherlands and was carried out in 2012 to synthesize and summarize findings and lessons learned from evaluations published in the three-year period 2009 to 2011. The second review was commissioned by the Evaluation Office of UNICEF and covered the period 2012 to 2015. This review summarizes the findings and lessons learned gleaned from a sample of evaluations published from the beginning of 2016 to the end of 2019.

The lessons from this review are expected inform on-going policy, strategy, and program development to best achieve results for children. All three reviews were carried out using a methodology and approached endorsed by the OECD DAC Network on Development Evaluation.¹

Approach and Methodology

The methodology used in this DER are consistent with the methods and approaches used in 2012 and 2016. This involved selecting a sample of published evaluations adjusted to be representative of UNICEF programming in different thematic areas of programming, over respective years of programming and by regional concentration of UNICEF activities (including humanitarian programming). The reports were then assessed to categorize findings across six evaluation criteria and 20 sub-criteria. A set of global thematic evaluations was the subject of a qualitative review.

For each criteria and sub-criteria, the review team classified findings on a four-point scale (highly unsatisfactory, satisfactory, unsatisfactory, and highly unsatisfactory) in accordance with detailed classification guidelines. In addition, the review team identified contributing factors which contributed to either a positive or a negative finding. The results of this review process are discussed in detail in Section 4.0.

Assessment Criteria

1. Relevance of the Intervention
2. The Achievement of Objectives
3. Cross Cutting Themes
4. Sustainability
5. Efficiency
6. Using Evaluation and Monitoring to Improve Effectiveness

Conclusions

Relevance

1. UNICEF has **consistently demonstrated a high level of relevance** in the programs and activities it supports over a ten-year period encompassing all three effectiveness reviews (2012, 2016 and 2020). UNICEF supported programs are **well suited** to the needs and

¹ UNICEF DER 2016

(https://www.unicef.org/evaldatabase/files/UNICEF_Review_of_Development_Effectiveness_2012_2015.pdf)

priorities of the children and adolescents they target, remain consistent with national priorities, and often provide benefits at a national scale. There is a **clear link from careful needs analysis during program design** to programs that are **well suited to the needs and priorities of children and adolescents**. UNICEF can strengthen this link by continuing/increasing priority and **emphasis given to consultation with** children, adolescents, and other **community stakeholders** during program development.

2. UNICEF has demonstrated a **strong commitment to engage in and support partnerships with a wide range of stakeholders** including national government agencies, civil society organizations, UN Country Team (UNCT) members and other bilateral and multilateral development partners. This represents an important **organizational comparative advantage**.
3. Building on strong partnerships with governments and other stakeholders, UNICEF has demonstrated a **capacity to support programming taken to scale at a national level**. This combined with the fact that programs suit the needs of children and adolescents provides evidence of a strong positive contribution to outcomes at a national level – including progress toward child-relevant SDGs.

Objectives Achievement

4. UNICEF has maintained a **strong record of positive program performance in objectives achievement** across all three effectiveness reviews. This finding is reinforced by the positive results of the 2020 Development Effectiveness Review. Almost all UNICEF supported programs and activities (95 percent) provide significant benefits to target group members. However, **there is scope to improve targeting** to provide more direct benefit from UNICEF supported programs to **marginalized** children and adolescents.
5. UNICEF has made an **effective contribution to the development of national policies, plans and programs for children** as measured across all three effectiveness reviews . In recent years (2016 to 2019), UNICEF evaluations have focused more closely on the operational effectiveness of programs than on “upstream” effects on policy. This represents a **shift in the focus of evaluations rather than a reduction in the level of policy advocacy by UNICEF**.

Cross-Cutting Themes: Gender Equity, Environment, Promoting Equity

6. UNICEF supported programs continue to improve in the level of attention and the positive results achieved to improve Gender Equity. This is reflective of **increased attention to Gender Equity under successive Gender Action Plans**. In addition, **UNICEF evaluations are addressing Gender Equity more consistently** and with more in-depth analysis. Further, the improvement in results for contributing to Gender Equity demonstrates an organizational ability to successfully implement strategies for change at UNICEF. Nonetheless there remains an opportunity to build on the steady progress on Gender Equity made during the 2016 to 2019 period and to raise **the overall level of organizational ambition with regard to gender equity** as recognized by the recent evaluation of the GAP.

7. UNICEF evaluations **do not generally address contributions to environmental sustainability in any systematic way**. The reviewed evaluation reports made **very little reference to efforts to address or mitigate the effects of climate change** in the programs and activities UNICEF supports. This **lack of attention** to the broader positive or negative environmental effects of UNICEF supported activities, including those related to climate changes, **weakens UNICEF's ability to develop and support a credible contribution story regarding climate change**: despite a strong organizational track record in, for example, addressing water supply insecurity in climate change affected regions and countries. **Evaluation and monitoring systems** and practices at UNICEF will **need to clearly address issues of environmental impact and climate change mitigation** to strengthen legitimate claims of UNICEF contribution.
8. The 2020 effectiveness review reports **positive results for UNICEF efforts to advance equity for marginalized and vulnerable children and adolescents** – most often with program elements targeting **adolescents and children with disabilities** – a reflection of recent organizational emphasis on these two groups of disadvantaged. This is another area where **increased emphasis on meaningful engagement and consultation with community members**, especially vulnerable children and adolescents and their families during program development **can improve UNICEF effectiveness**. This would assist UNICEF in **extending the application of the principles leave no-one behind** to a wider set of the most vulnerable children and adolescents.

Sustainability

9. The **sustainability of the benefits of UNICEF supported programs, especially financial sustainability remains a critical challenge**. Experience with programs that have proven sustainable indicates that responding more effectively to this challenge will require: intensified pre-program **planning for resource mobilization** from non-UNICEF sources on program completion; ensuring that **pilot, locally-scaled programs are designed with adequate plans for going to scale**, including cost effectiveness data gathering; and ensuring that UNICEF supported programs include realistic **exit strategies** for the period when UNICEF support is being withdrawn. While UNICEF has been generally successful in supporting programs delivered at scale, evaluation reports indicate their medium to longer term financial viability is a continuing challenge. There is also a need to **ensure that UNICEF supported programs are embedded in national budgets** with access to sustainable financing from **governments, the private sector and civil society**. SP Change Strategy 5 – Leveraging the Power of Business and Markets for Children may provide an important avenue to improving the sustainability of at-scale interventions and programs supported by UNICEF.
10. UNICEF has been **successful in developing the capacity of partner institutions and communities** so that they can continue with the delivery of services following program completion (provided financing can be secured). This reflects both a consistent investment in planning for capacity development and a long history of effective partnership at UNICEF.

Efficiency

11. **Efficiency and value-for-money issues are receiving more attention in recently published UNICEF evaluations.** In addition, the 2020 review indicates that **more UNICEF offices and staff are demonstrating a “cost-effectiveness mind set”** in developing and supporting programs. However, **cost and resource-efficiency remain a major challenge.** UNICEF supported programs can **also encounter considerable delay in implementation** due to over-ambitious timelines, administrative delays in approvals, problems in disbursement and issues in supply-chain management. These findings have been consistent across all three effectiveness reviews and have shown little to no improvement. The **current initiative to re-examine UNICEF business models** and operational practices in different country contexts (such as income category) will need to address questions relating to **improved business practices for timely and efficient program delivery.**

Using Monitoring and Evaluation to Improve Effectiveness

12. UNICEF **systems for evaluation have steadily improved in overall effectiveness over ten years of programming** and across all three effectiveness reviews. This improvement is supported by the concurrent improvement in the quality scores of evaluation reports assessed in the 2012, 2016, and 2020 effectiveness reviews. In addition, the number, responsiveness and **quality of the management response to evaluation reports have significantly improved.**
13. The **UNICEF Strategic Plan, 2018 -2021**, included a **comprehensive Results Framework with results, related performance targets, and indicators** developed at the Goal, Outcome and Output level. Wherever possible, especially at output level, the results framework also includes baseline data. **Over time this should improve evaluation findings related to the strength of monitoring and RBM systems** at UNICEF. However, evaluations reports reviewed in 2020 indicate that **the effectiveness of monitoring and RBM systems at country program level remains a critically important challenge.** Often the absence of a credible baseline and valid change indicators reported against time-bound targets frustrated efforts at results reporting and RBM. If UNICEF is to improve the visibility and credibility of claims regarding its contribution to changes at country level it will need to **better link the baselines, targets and results of the SP Results Framework to program monitoring systems in operation at country level.**

Innovation

14. UNICEF supported programs and activities reflect a **significant effort to support innovation.** Innovations **in program approach**, including the application of new models to service delivery **are more frequent** in UNICEF supported programs than “hard innovations” such as new information and communications technology. UNICEF has signaled its strategic intent to use innovation in the current SP but there is still considerable room **for more strategic and systematic use of innovation** in at least half of the programs covered by evaluations in the 2020 review. As demonstrated by the

relative success of the Gender Action Plans and as pointed out in the recently published thematic evaluation of innovation, **greater effectiveness through innovation** may require the development of “an explicit overarching **innovation strategy**”, a **better understanding of staff capacity** for innovation, **better communications** on innovation as a priority and **improved accountability** for innovation.

Evaluating the Strategic Plan 2018-2020

15. The Development Effectiveness Review 2020 provides **useful evidence** regarding **base-line conditions** prevailing at the beginning of the Strategic Plan for some of the most important **change strategies (CS) and organizational enablers (EN)**. In particular, the findings of the review support the thesis that **CS1: Programming at Scale** for Children represents an important comparative advantage for UNICEF based on the relevance and scale of UNICEF programming and its track record on effective partnerships. However, inadequate integration into national budgets for some programs suggest that **greater attention to securing national funding** could improve effectiveness in both going to scale and sustaining program benefits.
16. The results of the 2020 review indicate that **UNICEF is well positioned to capitalize on CS2: Gender-Responsive Programming**. The mainly successful implementation of two sequential Gender Action Plans (2014-17 and 2018-21) provides important evidence that UNICEF can **be successful by making significant investments in organizational architecture** (including staff capacities and programming guidelines) to achieve **changes at a strategic and operational level in pursuit of clear priorities**. It also indicates that there is continuing scope for increased UNICEF ambition and attention to Gender Equity which **further raises the importance of CS2**. This is an important lesson for the Strategic Plan. It suggests that continued investment of time, attention and resources to CS5 **Leveraging the Power of Business** can be an effective change strategy.
17. Regarding **CS 3: Winning support through stronger advocacy**, fundraising and communications for children, UNICEF has demonstrated a **strong track record of effectiveness in advocating for changed policies and programs** to the benefit of children, especially at national level. Evaluation reports, rarely address the question of fundraising, however, they do indicate that **securing national budgetary commitments is critical to sustaining the benefits of UNICEF supporting programs**.
18. UNICEF has contributed to a **long-term trend in improved results in sustainable capacity development with, in particular, national institutions of governance**. This should allow **UNICEF to build on CS4: Developing and leveraging resources and partnerships for children by increasing the leveraging of national resources**. UNICEF can also rely on demonstrated effectiveness in advocacy and the use of its convening and coordinating powers to provide scope for leveraging resources.
19. **CS5, Leveraging the power of business and markets, is a relatively new strategic priority** for UNICEF, introduced in the Strategic Plan 2018-2021. Not surprisingly, it has not been examined in the evaluation reports reviewed in 2020 since almost all were commissioned before 2019. The advent of programmatic action such as the Business 4

Results (B4R) initiative and continued senior management support represent actions taken since 2018 to act on CS5. There is **an opportunity for UNICEF to study recent experience** in engaging with businesses on common objectives to identify and reduce barriers to effective cooperation and to **identify good practices and refine guidelines** so that country offices can confidently implement CS 5.

20. UNICEF has been **effective in applying Change Strategy 6: United Nations working together** based mainly on a long-standing commitment to **partnership with UNCT members and national governments**. It has also demonstrated the capacity to implement Change Strategy 7 – Fostering Innovation. However, up to 2018 at least, UNICEF lacked a fully developed plan and strategy for supporting innovation.
21. UNICEF has **taken action to strengthen results reporting at the Strategic Plan level** with the establishment of the DAPM and its efforts to improve country office reporting. In addition, the 2018-2021 SP, was the first to roll-out a comprehensive Results Framework with detailed results definitions, indicators, baselines and targets at the goal, outcome and output levels. This should, over time, help UNICEF to act more effectively in implementing **CS8: Using the power of evidence to drive change for children** and **EN 2: Results oriented, effective and efficient management**. This would help UNICEF to strengthen the credibility of the “contribution story” it makes regarding results at the goal, outcome and output levels.
22. Building on the SP Results Framework to effectively realize **EN2: Results oriented management** will require UNICEF to: a) **better link baselines, targets, and results data gathered for SP reporting requirements to monitoring systems at country program level** and b) **better link results reported at the output level** (where UNICEF has more control and accountability) **to results at outcome and goal level** where contribution and accountability are shared among many different actors.

1. Introduction

1.1 Background and History of Effectiveness Reviews at UNICEF

This is the third consecutive development effectiveness review (DER) undertaken at UNICEF to synthesize findings from its own published evaluation reports. The first was commissioned by the Ministry of Foreign Affairs of the Netherlands and was carried out in 2012 to synthesize and summarize findings and lessons learned from evaluations published in the three-year period 2009 to 2011. The second review was commissioned by the Evaluation Office of UNICEF and covered the period 2012 to 2015. This review summarizes the findings and lessons learned gleaned from a sample of evaluations published from the beginning of 2016 to the end of 2019 and, in Section 4.0, compares results across all three reviews (2012, 2016, and 2020).

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. All three reviews were carried out using a methodology and approach endorsed by the OECD DAC Network on Development Evaluation.²

1.2 Purpose

The purpose of this DER is to generate a body of credible comparative information on a common set of criteria (as prescribed in the Guidance Document³ and as used in the two prior reviews) that provides an insight into the development effectiveness of UNICEF. The review involves conducting a systematic synthesis of information from available evaluations published in 2016-2019 as they relate to the criteria of development effectiveness.

The DER report is envisaged to be used in the following ways:⁴

- By UNICEF managers, to gain an independent external view of UNICEF's 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 Plan for Global Evaluation 2018 - 2021;⁵ and,

² UNICEF DER 2016

(https://www.unicef.org/evaldatabase/files/UNICEF_Review_of_Development_Effectiveness_2012_2015.pdf)

³ Assessing the Development Effectiveness of Multilateral Organizations: Guidance on the Methodological Approach (<http://www.oecd.org/dac/evaluation/dcdndep/50540172.pdf>)

⁴ UNICEF Development Effectiveness Review, 2016-2019 - Concept Note, December 2019

⁵ https://www.unicef.org/evaldatabase/files/UNICEF_Plan_for_Global_Evaluations_9_JAN_2018.pdf

- By stakeholders in UNICEF's Evaluation Office.

1.2.1 Input to Evaluation of the Strategic Plan

This Development Effectiveness Review was commissioned to coincide with the evaluation of the UNICEF Strategic Plan 2018 – 2021 which has as its stated purpose:

“Identify lessons learned from the contribution of the Strategic Plan 2018-2021 to the achievement of the 2030 agenda, prior to the onset of COVID-19.”⁶

As a result, the Evaluation Office and the authors of this report have sought to ensure that the review contributes as much as possible to the ongoing evaluation of the Strategic Plan while remaining true the methodology and analytical model used in the 2012 and 2016 reviews and, thereby, maintain a valid time series of findings over the ten years of evaluation results systematically synthesized in the three effectiveness reviews.

This raises the question of how the evaluation findings synthesized in this review can make a valid contribution to the evaluation of the Strategic Plan.

It is important to begin by noting that program evaluation is an inherently retrospective activity. Evaluations commissioned in one year are often not completed or published for a year or more and the programming covered by each evaluation will have been designed and approved some considerable time before the evaluation was undertaken. Because this review encompasses evaluations published in 2016, 2017, 2018 and 2019, the programming under evaluation was delivered by UNICEF (mainly through its partners) across two Strategic Plans (2014 -2017 and 2018 – 2021).

The sample of 62 evaluations analyzed for this review were chosen in a randomized process and adjusted to best represent UNICEF programming by thematic area of program activity, regional distribution of expenditure and year of publication. This was necessary to maintain consistency of methodology across the three reviews and preserve a time series of evaluation lessons which can be extended in future reviews. As a result, it is not possible to separate findings applicable to one or the other of the Strategic Plans covered by the review.

Rather, the evaluations findings and lessons learned presented here should be seen as representing the **transition from the period before the Sustainable Development Goals (SDGs) were agreed in 2016 to their full recognition in the UNICEF Strategic Plan 2018-2021.**

⁶ Draft Concept Note for the Evaluation of the UNICEF Strategic Plan 2018 - 2021

2. Applying Review Results to the SP Evaluation

2.1 Highlights of the Strategic Plan

The UNICEF Strategic Plan 2018 -2021 focuses on realizing the rights of every child,⁷ especially the most disadvantaged. The Strategic Plan builds on the work of the preceding plan, which in turn was informed by major global movements and discussions on the Sustainable Development Goals (SDGs) and the lessons learned and findings from the review of the UNICEF's Medium-Term Strategic Plan (MTSP) 2006-2013.⁸

The current Strategic Plan continued the process of aligning UNICEF operations to the SDGs and the achievement of the Agenda 2030. The Strategic Plan is anchored in the Convention on the Rights of the Child and intended to “help realize the SDG’s vision of a work in which no one is left behind”.⁹ It includes an overarching goal of realizing the rights of every child, especially the most

disadvantaged. The structure of the Strategic Plan drives toward this overall goal through 5 Goal Areas, 8 Change Strategies, and 4 internal enablers.

UNICEF Strategic Plan identifies eight change strategies and four organizational enablers to drive results.¹⁰

Box 1: Five Goal Areas

1. Every child survives and thrives.
2. Every child learns.
3. Every child is protected from violence and exploitation.
4. Every child lives in a safe and clean environment.
5. Every child has an equitable chance in life.

UNICEF Strategic Plan 2018 - 2021

Box 2: Change Strategies and Organizational Enablers

Change Strategies (8)

- Programming excellence for at-scale results for children.
- Gender-responsive programming.
- Winning support for the cause of children from decision-makers and the wider public.
- Developing and leveraging resources and partnerships for children.
- Leveraging the power of business and markets for children.
- United Nations working together.
- Fostering innovation in programming and advocacy for children.
- Using the power of evidence to drive change for children.

Organizational Enablers (4)

- Responsive, transparent and accountable internal governance.

⁷ UNICEF mandate covers all children below the age of 18 years, as the Convention on the Rights of the Child.

⁸ 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.

⁹ UNICEF, Strategic Plan 2018-2021, p.9.

¹⁰ UNICEF Strategic Plan 2018 - 2021

- Results-oriented, efficient and effective management.
- Staff capacity to drive change for children.
- Versatile, safe and secure knowledge and information systems.

UNICEF Strategic Plan 2018 - 2021

2.2 A Focus on Change Strategies and Organizational Enablers

The “Report on the midterm review of the UNICEF Strategic Plan, 2018–2021 and annual report for 2019 of the Executive Director of UNICEF”¹¹ along with supporting documents presents a structured, quantified assessment of how UNICEF has contributed to progress toward the SDGs and Agenda 2030. These documents represent the best available data on results achieved under the Strategic Plan. With this material as a base it seems clear that the best use of this Development Effectiveness Review as a component of the Strategic Plan Evaluation can be found in the relationship between evaluation findings and the Plan’s critical change strategies and organizational enablers.

These change strategies and organizational enablers can be viewed as an important lever to allow the Strategic Plan contribute to the achievement of the SDGs and the realization of Agenda 2020. They represent an effort **to capitalize on organizational comparative advantages, address ongoing challenges and marshal resources** to effectively support the results targets identified under each of the five goal areas. In this regard, the evaluation findings and lessons learned documented here can best be seen as an exploration of how the change strategies and enablers build on strengths identified in evaluations from 2016 to 2019 or, alternatively, represent an effort to strengthen areas identified as challenges.

Section 5.0 discusses the findings of the DER in light of the ongoing evaluation of the Strategic Plan and the implications of those findings for changes to the Plan going forward.

3. Methodology and Assessment Criteria

3.1 Methodology: A Brief Overview

¹¹ UNICEF, “Report on the midterm review of the UNICEF Strategic Plan, 2018–2021 and annual report for 2019 of the Executive Director of UNICEF”, May 2020, p.1

As already noted, the methodology and assessment criteria used in this DER are entirely consistent with the methods and approaches used in 2012 and 2016. This involved selecting a modified random sample of published evaluations adjusted to be representative of UNICEF programming in different thematic areas of programming, over respective years of programming

Box 3: 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

and by regional concentration of UNICEF activities (including humanitarian programming). Each evaluation in the sample was then reviewed from a quality assurance

perspective. The reports (and accompanying Management Responses) were then assessed to categorize findings across 6 main evaluation criteria and 20 sub-criteria. In addition, a set of global thematic evaluations was the subject of a qualitative review to complement the results of the quantitative assessment of findings.

For each criteria and sub-criteria, the review team classified findings on a four-point scale (highly unsatisfactory, satisfactory, unsatisfactory, and highly unsatisfactory) in accordance with detailed classification guidelines. In addition, the review team identified contributing factors which contributed to either a positive or a negative finding. The results of this review process are discussed in Section 3.0 below.

The sample selection process resulted in a set of 62 evaluation reports for analysis. This is consistent with practice in 2016 (66 evaluation reports in the ultimate sample) and 2012 (62 evaluation reports). For a full discussion of the methodology used in this DER, including a comparison of the sample and the universe of available evaluation reports see Annex C.

3.2 Evaluation Coverage

A key methodological question for any review of this type is how well or poorly specific evaluation criteria are covered in the evaluation reports under review. The review team has addressed this by categorizing the level of coverage of a given criteria based on the number of evaluation reports which have provided valid, verifiable evidence to support a specific finding for a given criteria. This allows the review to arrive at the number (n) of valid observations for each of the sub-criteria under review. Where n is 46 or higher, coverage is classified as strong; moderate coverage means n is between 31 and 45 and weak coverage is assigned when less than 31 (or half) of reviewed evaluations provide a response the criteria in question. Table 1 illustrates the level of coverage attained in each of the three effectiveness reviews.

Table 1: Level of Coverage for Each Evaluation Criteria

Criteria and Sub-Criteria	n* 2020	Coverage Level 2020†	n* 2016	Coverage Level 2016†	n* 2013	Coverage Level 2012†
Relevance						
1.1 Programs suited to the needs and/or priorities of the target group.	55	Strong	56	Strong	60	Strong
1.2 Programs align with national development goals.	50	Strong	56	Strong	61	Strong
1.3 Effective partnerships with governments, bilateral and multilateral development and humanitarian organizations and NGOs.	54	Strong	46	Strong	56	Strong
Achievement of Objectives						
2.1 Programs and projects achieve their stated humanitarian and development objectives.	49	Strong	63	Strong	61	Strong
2.2 Programs and projects have resulted in positive benefits for target group members.	60	Strong	64	Strong	62	Strong
2.3 Programs and projects made differences for a substantial number of beneficiaries.	54	Strong	58	Strong	61	Strong
2.4 Programs contributed to significant changes in national humanitarian and development policies and programs.	42	Moderate	55	Strong	43	Moderate
Cross-Cutting Themes						
3.1 Extent UNICEF support addresses Gender Equity.	50	Strong	58	Strong	42	Moderate
3.2 Extent to which changes are environmentally sustainable.	5	Weak	12	Weak	15	Weak
3.3 Programs effectively address issues of equity of children and adolescents.	56	Strong	54	Strong	N/A	N/A
Sustainability						
4.1 Benefits likely to continue after project or program completion.	57	Strong	56	Strong	59	Strong
4.2 Projects and programs are reported as sustainable for institutional and/or community capacity.	56	Strong	56	Strong	59	Strong
4.3 Programming contributes to strengthening the enabling environment for development.	48	Strong	48	Strong	40	Moderate
Efficiency						
5.1 Program activities are evaluated as cost/resource efficient.	53	Strong	52	Strong	48	Strong
5.2 Implementation and objectives achieved on time	41	Moderate	39	Moderate	39	Moderate
5.3 Systems and procedures for project/program implementation and follow up are efficient.	43	Moderate	42	Moderate	35	Moderate
Evaluation and Monitoring						
6.1 Systems and process for evaluation are effective.	52	Strong	56	Strong	42	Moderate
6.2 Systems and processes for monitoring and reporting are effective.	59	Strong	60	Strong	59	Strong
6.3 Results based management systems are effective.	58	Strong	61	Strong	24	Weak
6.4 Evaluation is used to improve humanitarian and development effectiveness.	62	Strong	66	Strong	61	Strong

‡ n= number of evaluations addressing the given sub-criterion

† Strong: n = 46 and above; Moderate: n = 31 – 45; Weak: n = less than 31

Table 1 clearly illustrates the consistently high level of coverage afforded all six main criteria and the 20 sub-criteria used to assess program effectiveness across the three reviews (2020, 2016 and 2012). In the 2020 sample, 16 criteria received coverage rated as strong, while three were rated moderate and only was weak (criteria 3.2 on environmental sustainability which was covered in only seven evaluation reports). This coverage pattern is consistent with both prior reviews. As a result, Section 4.0 on findings provides results for 19 all six main criteria and 19 of the 20 sub-criteria in table 1.

3.3 Limitations

As a meta-synthesis of previously completed evaluation reports, this development effectiveness review is 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 the 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 2016 to 2019. 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:** 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 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).
- **The Retrospective Nature of Meta-Synthesis:** Evaluation reports are, by definition, retrospective in that they provide ex-ante assessments of the effectiveness of programs.

While these limitations are real, none strongly undermine the utility of the results of the development effectiveness review process. The review provides a meaningful overview of the existing evaluative evidence of organizational effectiveness (and how that may have 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 already been made by organizing and analyzing the results of those investments in a systematic way.

4. Evaluation Findings

This Chapter presents the results of the review as they relate to the six main criteria and associated sub-criteria used to assess effectiveness. 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 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 positive findings. Trends in results over time and differences among, for example, country income categories for each evaluation sub-criterion were identified by using a system of weighted average scores with a numeric value assigned to each findings classification – from highly satisfactory to highly unsatisfactory (see Annex C, Methodology).

Assessing and reporting on contributing factors

In reporting on the frequency with which contributing factors are cited the report (as in 2012 and 2016) 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. When

reporting on either a positive or negative factor influencing results, the report includes in brackets some of the countries where evaluations referred to a specific factor. This list of countries and evaluations is not exhaustive and is intended only to highlight the most important examples.

Box 4: 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

4.1 Relevance

Summary

The relevance of UNICEF supported programs and activities is a major and consistent strength of UNICEF. These programs are highly suited to the needs and priorities of target group members – often due to effective consultations with and participation by stakeholders during the design program design phase. In those instances where programs are not well-suited to the needs and priorities of the target group this most often stems from weak or missing research on needs analysis or insufficient consultation, especially with community groups and representatives of affected children, adolescents and their families.

Of equal importance, UNICEF support has been well aligned with the national strategies, priorities and programs across all three effectiveness reviews (2012, 2016, 2020) and throughout the past and current Strategic Plans. This represents an important comparative advantage for UNICEF. Positive alignment with national goals as measured by evaluations reviewed in 2012, 2016, and 2020 was most often arrived at by:

- Ensuring UNICEF supported programs and activities contribute to the goals and targets of national plans and priorities at a sector and sub-sector level, including national action plans for addressing key issues for children and adolescents;
- Engaging with key stakeholders at national and sub-national levels to ensure that supported programs and activities are coordinated with existing and planned programs; and,
- Ensuring active UNICEF participation in efforts to develop a common UNCT approach through mechanisms such as the UN Partnership Framework (UNPAF) agreements.

Finally, UNICEF has demonstrated a strong capacity to engage in and support effective partnerships within the United Nations family and with national government agencies, national NGOs, international NGO, bilateral development partners and civil society organizations (CSOs). A particularly important feature of partnership as practiced by UNICEF was a willingness to use informal structures to engage with community members and to link these structures to more formal government-led systems.

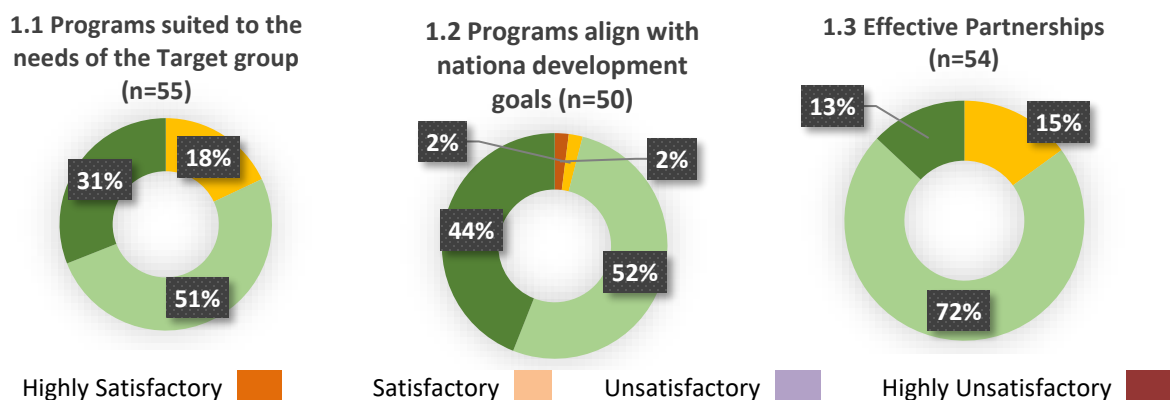
Coverage

As in the 2012 and 2016 reviews, all three sub-criteria focused on relevance were rated strong for coverage with 50 or more evaluation reports in the sample presenting findings relevant to each of the criteria (see figure 1 below).

Key Findings: Relevance

Relevance remains a key strength of UNICEF programming as reported in the evaluations reviewed. For all three relevance sub-criteria, findings were either satisfactory or highly satisfactory in over 80 percent of the evaluation reports reviewed.

Figure 1: Results for Relevance Criteria



Evaluation reports have consistently shown that UNICEF is **adept at aligning programming with national development goals and priorities** as expressed in national strategies and programs (sub-criteria 3.2) , including for national targets for achieving the SDGs. The 2020 result, with 96 percent of evaluations reporting positive findings for this sub-criteria, is a continuation of the high levels reported in 2012 and 2016 which both reported positive findings in more than 95 percent of evaluations.

The strength of UNICEF performance in aligning with national priorities is directly related to its track record for **developing and sustaining effective partnerships** with national government partners, bilateral and multilateral development partners, humanitarian agencies and NGOs. Effective partnership can be seen **as an important area of comparative advantage for UNICEF in the 2016 to 2019 period**. In fact, performance in this area has improved when compared to the 2016 review. This review found that 85 percent of evaluations reported satisfactory or highly satisfactory findings regarding the strength of UNICEF partnerships compared to 70 percent in the 2016 review. This represents a return to the high level of performance reported in 2012. One element in UNICEF’s success at partnership had been an ability to serve as a **convening and coordinating platform for partners**. UNICEF has been very active in supporting and participating in coordinating mechanisms at both national and sub-national levels.

Box 5: Effective Use of UNICEF Convening Power
 UNICEF was active in developing and supporting coordination mechanisms for the roll out of a Life Skills Education (LSE) Model across schools at state level in India. UNICEF also supported the work of the Child Protection Working Group (CPWG) and the SGBV Task Force in Lebanon which contributed to improvements in child protection and GBV programming across the country.

Significantly, the reviewed evaluation reports consistently find that UNICEF supported programs and activities are **suited to the needs and priorities of the targeted beneficiaries** (Sub-Criteria 1.1). A full 82 percent of evaluation reports find that UNICEF programs and activities are well suited to the needs and priorities of target group members to either the

satisfactory or highly satisfactory level. This represents an improvement from the 70 percent level reported in the 2016 review, but not yet a return to the level of 90 percent found in the 2012 review.

The reviewed evaluation reports provide many examples of how UNICEF supported interventions match the needs and priorities of target group members across all five goal areas.

Table 2: Examples of Meeting Target Group Needs

Goal Area	Selected Examples
Survive and thrive	<ul style="list-style-type: none"> • Nutritional and hygiene counselling reported as relevant by target group members (Kenya) • Community health component tailored to respond to the needs of the most vulnerable population (Guinea Bissau) • Implementation modalities relevant to the needs of pastoralist communities (Ethiopia)

Goal Area	Selected Examples
Learn	<ul style="list-style-type: none"> Life skills education (LSE) tailored to identified needs of adolescent boys and girls (India)
Protection from violence and exploitation	<ul style="list-style-type: none"> Most adolescent girls and women rated GBV activities as relevant (Lebanon) Skills development programs relevant to needs of staff of courts and police (Kyrgyzstan)
Safe and clean environment	<ul style="list-style-type: none"> Earlier studies along with a diagnostic and needs assessment of health centers to identify needs related to water, latrines and waste management (Mali)
Equitable chance in life	<ul style="list-style-type: none"> Needs of the targeted beneficiaries could best be met through the cash transfer mechanism (Nigeria)
Multiple goal areas	<ul style="list-style-type: none"> Communications for Development (C4D) initiatives using messages and methods relevant to the community context (Bangladesh)

Some evaluation reports noted that the the high level of fit between UNICEF supported programs and activities and the needs of target group members is often grounded in a high level of dialogue with communities and their members (Box 6).

Box 6: Community Consultation to Identify the Needs of Roma Children

- In Albania dialogue between community and local institution was established for a better identification of problems that Roma children face.
- In Macedonia, implementing partners and other relevant Roma NGOs organized a series of workshops, door-to-door visits and informal meetings with Roma parents in all relevant communities to identify their needs, problems and suggestions for the intervention.
- In Serbia, a workshop at the beginning of the project gathered relevant partners and the local expert community who identified the pressing needs of marginalized children as well as ways of action to address them. In addition, an in-depth study with Roma families was conducted to identify early childhood development risks and vulnerabilities of Roma children. Roma CSOs were actively involved in project design, which brought their knowledge and experience to this process.

Source: Evaluation of “Breaking the cycle of exclusion for Roma children through Early Childhood Development and Education” multi-country project in Macedonia, Serbia and Albania, 2017, p.36-38.

Changes Over Time

Evaluations published by UNICEF have reported consistently positive results for the relevance of UNICEF programming across all three Development Effectiveness Reviews (Annex K).

Contributing Factors: Relevance

Sub-criterion 1.1: Programs suited to the needs of the target group.

The single most important factor supporting programming which is well suited to the needs of target group members is a formal process of needs assessment, baseline study or detailed context

analysis undertaken during the development and design phase. This was cited in more than half of relevant evaluations.

Table 3: Examples of Needs Assessment Strategies

Goal Area	Selected Examples
Survive and thrive	<ul style="list-style-type: none"> • Joint rapid assessment of needs for nutrition intervention in Vietnam
Learn	<ul style="list-style-type: none"> • Joint meetings with parents, communities and primary school directors (Argentina) • Community consultations to identify priorities and approaches to ECD and meeting learning needs of Roma Children (Albania, Macedonia, Serbia)
Protection from violence and exploitation	<ul style="list-style-type: none"> • Child protection and GBV needs identified through consultations and programming with refugees used to extend services to host communities (Lebanon) • Consultation with stakeholders to ensure interventions were relevant to operation of courts and police (Kyrgyzstan)
Safe and clean environment	<ul style="list-style-type: none"> • Diagnostic and needs assessment of health centers to identify needs related to water, latrines and waste management – included close consultation with (Mali)
Equitable chance in life	<ul style="list-style-type: none"> • Informal actors consulted on the use of community structures to prevent and respond to child welfare issues (Ghana) • Community consultations to identify needs and priorities of targeted beneficiaries of Cash Transfer Program (Nigeria)
Multiple goal areas	<ul style="list-style-type: none"> • See below:

In addition, UNICEF engages frequently in cross-sectoral, joint needs assessment activities during the program development process. Specific examples include a joint mapping episode with UNHCR and the national government (Lebanon), a comprehensive situation analysis of the needs of children and adolescents (South Sudan) and a joint refugee needs assessment mission by UNICEF/WFP/UNHCR in Congo.

One issue that was raised with some frequency in the evaluation reports reviewed was the need for UNICEF to be willing to support cross-sectoral work and to involve more partners in some programming. As noted in box 7, some evaluation reports show important evidence UNICEF using a multi-sectoral approach for program development. Participatory, cross sectoral needs assessment can be seen as a crucial factor in ensuring that program outputs are well suited to the needs of children and adolescents and their families, needs which may not be confined to a

Box 7: Multi-Sectoral Needs Assessment in Nigeria

The participation of UNICEF’s multi-sectoral team in the joint rapid assessment was a critical factor in formulating a framework with adequate focus on nutrition and WASH that have both immediate and long-term consequences for children. Priority humanitarian needs identified in the Multi-Sector Needs Assessment included the provision of clean drinking water, ensuring access to clean water sources, containing rising cases of diseases and addressing food shortages and associated malnutrition. The inclusion of activities responding to issues identified in the assessment including malnutrition, poor hygiene practices and inadequate WASH services in schools was welcomed by stakeholders

Source: Impact Evaluation of UNICEF Nigeria Girls’ Education Project Phase 3 (GEP3) Cash Transfer Program (CTP) in Niger and Sokoto States, 2017 (p85)

particular sector or sub-sector.

Consultations with and the engagement of community members (Box 6) in the development and design of interventions was also frequently cited as a positive factor in matching program outputs to client needs. Examples include consultations with community leaders before establishing family centres in Palestine, direct consultations with parents in Benin and joint planning with community members in Mali. Finally, a few evaluations pointed to use of existing or new community level data (sometimes from baselines studies) to effectively target poor and vulnerable households and individuals as in Indonesia, Eswatini, and Haiti.

Sub-Criterion 1.2: Programs Align with National Development Priorities

Not surprisingly almost all relevant evaluations (48 of 50) noted that UNICEF support to programs and activities was **designed to align with and referred specifically to national development plans and priorities** as expressed either in overall national development plans, sector plans, or joint priorities of national governments and the UN Country Team (UNCT). More precisely: 25 evaluations pointed to UNICEF efforts to ensure that programs aligned well with overall national development plans and programs; and 19 evaluations noted that UNICEF programs were well aligned with sector-wide or sub-sector plans and strategies.

Table 4: Examples of Alignment

Goal Area	Selected Examples of Alignment
Survive and thrive	<ul style="list-style-type: none"> Alignment with national health strategy, national health plan, vision statement or road map (Pakistan, Lebanon, Serbia, Zimbabwe, Jordan, Rwanda, Nigeria, Guinea Bissau, Serbia) Aligned with the national HIV/AIDS framework (Eswatini, Congo) Aligned with the National Child Survival Strategy (Zimbabwe)

Learn	<ul style="list-style-type: none"> Aligned to national plan/strategy for ECS and education for children and adolescents (Albania, Argentina, Benin, Burundi, Eswatini, India, Nigeria, Eswatini, Ethiopia, Macedonia, Myanmar, Romania, Serbia, Vietnam, Zimbabwe)
Protection from violence and exploitation	<ul style="list-style-type: none"> Aligned with the national plan for enhanced operation of law enforcement (Kazakhstan) Aligned with the National Strategy for the Rights of Children (Croatia) Aligned with the Child Protection and Welfare Act (Eswatini)
Safe and clean environment	<ul style="list-style-type: none"> Alignment with the national strategy for WASH (Cameroon, India, Lebanon, Sudan)
Equitable chance in life	<ul style="list-style-type: none"> Aligned with strategy/plan for poverty reduction (Haiti, Bangladesh)
Multiple goal areas	<ul style="list-style-type: none"> See below:

Finally, many evaluations highlighted the care taken to ensure that UNICEF supported programs and activities were aligned with **United Nations conventions, treaties, and country-specific UN agreements**. These include compliance with the UNDAF/UNPAF framework, alignment with the Sustainable Development Goals (SDG) and conforming with the requirements of international treaties including the United Nations Convention on the Rights of the Child (UNCRC) and the Convention on the Elimination of all form of Discrimination Against Women (CEDAW).

Sub-Criterion 1.3: Effective Partnerships

It is striking that most evaluation reports reviewed found that effective partnerships were a key element for UNICEF in effectively supporting programs and activities for children and adolescents. The most important positive factor noted (in more than half of reviewed evaluations) was an established history of UNICEF partnerships at country level including different groupings and purposes including but not limited to:

- Partnerships with ministries and agencies of national governments;
- Partnerships with international non-governmental (INGO) and civil society organizations (CSOs);
- Partnerships with national NGOS and CSOS;
- Partnerships at state and local level with government ministries and agencies;
- Partnerships with bilateral development agencies working in program countries; and,
- Partnerships with UN Country Team (UNCT) members and other multilateral organizations.

In many instances, the evaluations pointed out that UNICEF had helped to **convene a partnership** which crossed the groups listed above and allowed for multiple levels of partnerships.

Box 8: Coordinating Partners in Response to the Gorkha Earthquake in Nepal

UNICEF was uniquely positioned to support national coordination efforts, given its knowledge of local institutions and the confidence and working relations established with line ministries before the earthquake. UNICEF established an effective coordination working group, providing a coordination platform for partners and stakeholders addressing communication with beneficiary-related activities. At sub-national level the cluster coordination was largely delegated to INGO partners, and the engagement of local authorities was clearly a determinant in cluster coordination.

Source: Evaluation of UNICEF's Response and Recovery Efforts to the Gorkha Earthquake in Nepal,

Another factor in support of effective partnership by UNICEF is the organization's capacity to play an important role in **coordination**. Some evaluations point to UNICEF taking on a strong role in different mechanisms for both policy and operational coordination. A few examples include:

- Good communications for coordination with partners working at State level (Congo);
- Implementing a steering committee to coordinate institutional reforms (Kazakhstan);
- Developing standard procedures for case management (Jordan); and,
- Participating in working groups on child protection and Sexual and Gender-Based Violence (SGBV) (Lebanon).

Next in order of frequency cited was **the leadership, commitment and technical knowledge** demonstrated by UNICEF, often by establishing or leading coordination mechanisms (Pakistan, Sudan, Nepal, Nigeria).

4.2 Achievement of Objectives

Summary

UNICEF has established a strong record of objectives achievement over ten years of programming, including for this DER (2016 to 2019). More than three quarters of UNICEF programs and activities achieve their stated objectives and contribute to expected results. In the current review, almost all UNICEF supported programs (95 percent) provided significant benefits to target group members and 78 percent of evaluation reports found that the same programs had significant reach and contributed to national development goals.

One of the most significant findings of the 2020 review arises from the conjoining of results on two of the four sub-criteria. The fact that the overwhelming majority of evaluations found that UNICEF supported programs provide material benefits meeting the needs of the target group can

be linked to the finding that three quarters of these program provide benefits on a national or near-national scale. This suggests that UNICEF supported programs are contributing to results at national level and have the potential to make a significant contribution to outcomes for children and adolescents, including achievement of relevant SDGs.

In addition, UNICEF has contributed to the development of national policies and plans for children and adolescents during all three review periods. Findings for developing national policies are somewhat less positive for the 2020 review (74 percent) when compared to 2016 (82 percent). This may reflect a stronger focus on improving operational level performance by UNICEF and its partners. It may also reflect a set of evaluations more closely focused on implementation issues than on upstream efforts to address national policies. The reviewed evaluations highlight the factors that are most crucial to achieving positive results in the attainment of stated objectives:

- The requirement for focused, thorough and specific needs analysis which establish clearly which inequities or inadequacies in services or behaviour are being addressed (and what goals are being set for the program);
- The need for widest possible consultation with stakeholders, especially children, adolescents, their families and other community members (as well as key organizational stakeholders); and,
- A high level of flexibility in program design so that key elements can be adjusted to reflect changing contexts and new barriers to equity, access or positive behaviour change.

Coverage: Achievement of Objectives

Three of the four sub-criteria relating to objectives achievement received a strong level of coverage in the evaluation reports under review (2.1, 2.2, and 2.3). This pattern was consistent with the reviews undertaken in 2012 and 2016. Sub-criterion 2.4 regarding UNICEF's contribution to changing national policies and programs was covered to the moderate level with 42 of the 62 evaluations in the sample providing clear findings.

Key Findings: Achievement of Objectives

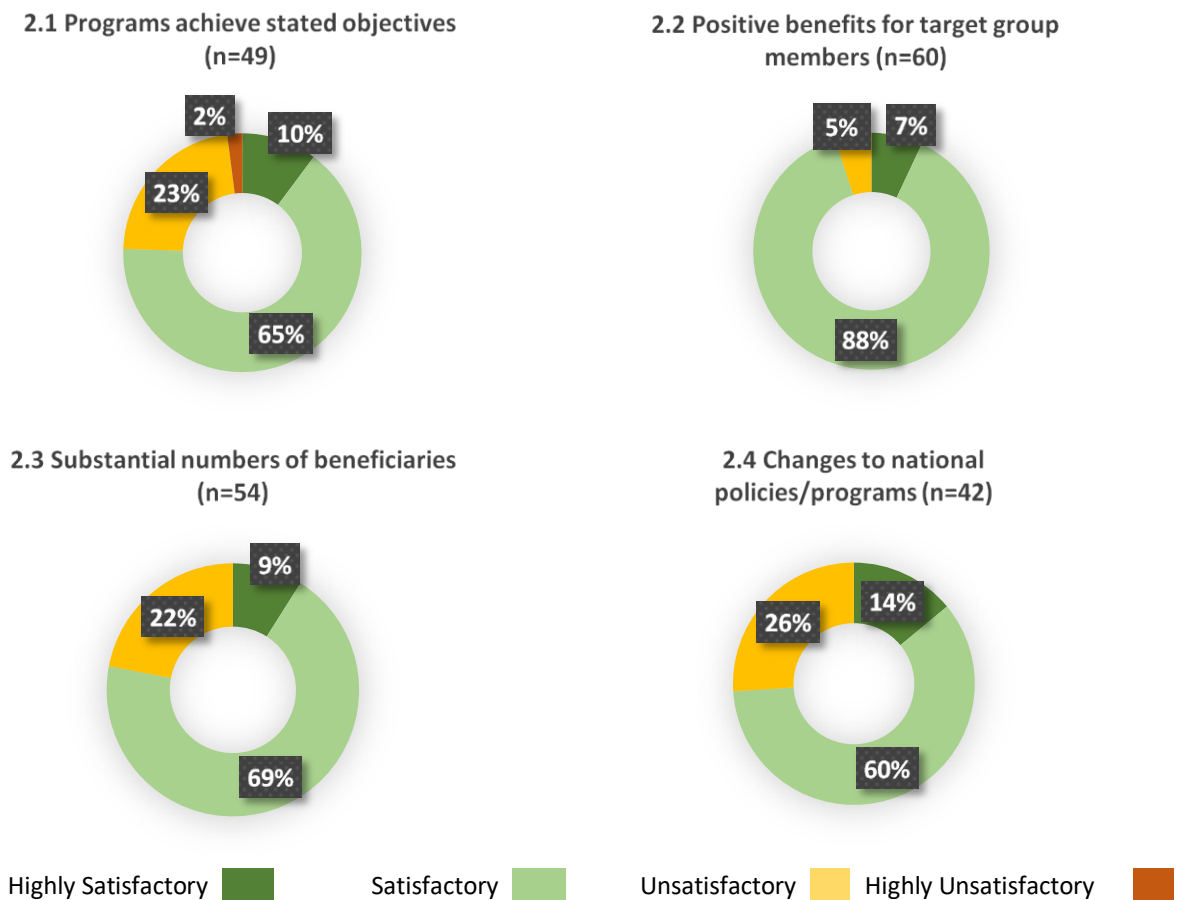
The 2020 DER found that 75 percent of the evaluations reviewed indicated that UNICEF supported programs and activities had achieved or exceeded their stated objectives. This was totally consistent with the results of the 2012 and 2016 reviews (77 and 75 percent respectively). Further, there was no indication in the evaluations reviewed that the targets or objectives set for UNICEF support were easily reached or insufficient in scope and ambition.

Results were positive across all four sub-criterion and especially so for 2.2 regarding whether or not UNICEF supported programs resulted in tangible benefits for target group members. For this area, 95 percent of evaluations reported either satisfactory or highly satisfactory findings. It is also important to consider the scale of programs and activities supported by UNICEF by examining whether or not evaluations identified substantial numbers of beneficiaries. If UNICEF support was overly focused on a small group of beneficiaries this criterion would likely be subject to mostly negative findings. **In fact, 78 percent of the evaluations reviewed found that**

UNICEF supported programs made a difference for substantial numbers of people and contributed to national development goals.

While fewer evaluation reports produced findings regarding UNICEF contributing to changes in national policies and programs (42 of 62), the overall results were positive with 74 percent reporting findings at the satisfactory or highly satisfactory level. It is also worth noting that this sub-criterion was rated highly satisfactory in 14 percent of the evaluations reviewed, the largest proportion of highly satisfactory rating for all four sub-criteria relating to objectives achievement. This may reflect especially effective efforts at advocacy and the provision of technical support to the development of national policies and programs critical to achieving results for children.

Figure 2: Results for Objectives Achievement Criteria



Examples of Objectives Achievement

Box 9: Achieving Objectives for Improving Critical Knowledge and Skills for Youth in Egypt

The Meshwary Project left an impact on the beneficiaries because it equipped these young people with critical knowledge and skills, which in turn enabled them to become socially empowered. In fact, more than 80% of the sampled beneficiaries indicated that the project was their main source of information about the life skills, career qualification skills, planning and feasibility and financial skills. Another key aspect to the project’s success relates to the assurance and sense of empowerment the beneficiaries had after working with Meshwary.

Source: Evaluation of Meshwary Project Phase II in Egypt, 2017 (p11)

It is worth considering more precisely what types of objectives are being achieved when UNICEF supported programs and activities are found to produce satisfactory or highly satisfactory results for the different sub-criteria of objectives achievement. Relevant examples of the benefits resulting from UNICEF support can be found across all five UNICEF goal areas.

Table 5: Examples of Objectives Achievement by UNICEF Goal Area

Goal Area	Selected Examples
Survive and thrive	<ul style="list-style-type: none"> • Reductions in toxic stress and relief from stress for children, parents and caregivers as a result of play-based learning (Uganda). • Increases in use of antenatal care and safe delivery contributing to improved maternal mortality (Algeria, Niger, Rwanda). • Increased use of Primary Health Care (PHC) services across a diverse set of facilities (Lebanon). • Positive changes in perinatal care and organizational culture associated with safe practices and a rights-based approach to services for women, newborn children and families in the health sector (Paraguay).
Learn	<ul style="list-style-type: none"> • Improved school enrolment and retention and reduced over-aging and repetition among both boys and girls (Benin, Burundi, Argentina, Kenya).
Protection from violence and exploitation	<ul style="list-style-type: none"> • Reducing or mitigating high risk behaviour (Jordan). • Encouraging fair treatment of juvenile offenders through a shift away from punitive measures to assistance as the aim of juvenile justice (Kyrgyzstan, Kazakhstan, Fiji). • Creating safe school environments (Eswatini). • Empowering girls through child protection activities which reduce inter-child violence and GBV (Lebanon).
Safe and clean environment	<ul style="list-style-type: none"> • Improved sanitation and hygiene conditions and practices including in refugee situations (Lebanon, Congo). • Improvements in the epidemiological profile among refugee populations as a result of access to clean water (Congo).
Equitable chance in life	<ul style="list-style-type: none"> • Increased awareness among policy makers, key decision makers and the public on issues and inequities affecting children along with increased political will to take action to address inequities (Thailand, Cameroon, Palestine).
Multiple goal areas	<ul style="list-style-type: none"> • Positive changes in the behaviour of children and adolescents as a result of play therapy (Honduras).

These are just some of the varieties of positive effects achieved when UNICEF supported programs are able to achieve their objectives, further emphasizing the importance of a satisfactory or highly satisfactory finding in 75 percent of the evaluations reviewed.

Changes Over Time

As with relevance criteria, those for objectives achievement show a stable pattern of positive results across the three effectiveness reviews. One slight exception is sub-criterion 2.4 which saw a slight decline in weighted average score from 2016 to 2020 (from 3.02 to 2.88). This may reflect a slight change in focus for evaluations published between 2016 and 2019 compared to those profiled in the earlier study. The analysts for the 2020 review found that evaluations were most often focused on objectives achievement in operational terms – did the programs provide the outputs on time and to the appropriate beneficiaries rather than how well they may have influenced national policies (as in sub-criterion 2.4). This may in-turn reflect the mix of programming chosen for evaluation in the most recent period rather than any decline in UNICEF performance regarding advocacy for changes in national policy.

Contributing Factors: Achievement of Objectives

Sub-Criterion 2.1: Achieving Stated Objectives

For the most part, when addressing why UNICEF supported programs and activities reach (or fail to reach) their stated objectives, evaluation reports refer the same set of positive and negative factors as listed for sub-criteria 1.1 (meeting the needs and priorities of the target group) and 1.2 (alignment with national priorities). The most important of these factors, referenced in most evaluation reports are: i) a formal process of needs assessment, baseline study or detailed context analysis undertaken during the development and design phase; and, ii) consultations with and engagement of community members in the development and design of interventions. Thus, positive results in achieving objectives are often as dependent on effective needs assessment and strong community consultation were positive findings on the relevance of UNICEF support.

Some evaluations pointed to the strength of partnerships with government agencies and NGOs advocating for child rights (Malawi) while others noted the strength of inter-agency and inter-personal communications (Bangladesh). When results for objectives achievement are less positive, this is often associated with problems in design during the development phase. In some cases, problems arise because program designs are too inflexible and are not adjusted as circumstances change. A few evaluations also pointed to the need for UNICEF and its partners to take a more multi-sectoral approach to program design so that more key stakeholders are involved in defining programs and finding solutions, however this was not linked to program performance in meeting stated objectives.

Sub-Criterion 2.2 Positive Benefits for Target Group Members

An important factor in ensuring that UNICEF supports programs which benefit target group members, cited in over just over half of the evaluation reports reviewed, is the extent of consultation with key stakeholders, including members of the target group. Sometimes this results in the use of specialized arrangements or methods for service delivery in order to access targeted beneficiaries, including providing workshops for parents to improve children's access (Croatia) and the development of mobile teams to provide health care services to refugees and ethnic minorities (Ethiopia).

Box 10: Reaching Ethnic Minorities with Mobile Health and Nutrition Teams in Ethiopia

Regular availability of services once a week is one of the strengths of Mobile Health and Nutrition Teams (MHNTs) compared to static facilities that are affected by the frequent breakdowns of services mainly due to staff turnover and stock out of supplies. In comparison, the MHNTs are reported to maintain the service provision at an acceptable level. One of the team leaders of MHNTs in Somali region who served for three years in different places stated that in the MHNTs 'never seen any stock out of medical supplies. Stakeholders (RHBS, health workers, implementing partners and beneficiaries) all agree that MHNTs in both regions have reached population groups (kebeles and woredas) that have not been beneficiaries of the static health systems.

Source: Evaluation of Mobile health teams in Afar and Somali Regions, in Ethiopia, 2016 (p24)

Criterion 2.3 Substantial Numbers of Beneficiaries – Program Reach

It is notable how often evaluation reports provide data on the significant scope of UNICEF supported programs and the numbers of beneficiaries reached. **A full 78 percent of evaluation reports provided data on program reach.** The most common factors cited for enabling programs to scale results include:

- Providing support to methods and approaches which are taken to scale by national ministries of, for example, health, education and justice (Pakistan, Zimbabwe, Kazakhstan);
- Partnering with national research institutes to provide systematic data collection on program reach (BRAC University in Bangladesh); and,
- Providing essential support to supply chains for essential health services, especially immunization (Zimbabwe).

In general, UNICEF supported programs reviewed in 2020 are less likely to include small scale, pilot projects which are not linked to plans for larger scale programming. Moving away from a pilot-focused approach may indicate that UNICEF supported programs are innovating more around methods for large-scale program delivery than around new technologies or methods of engagement (see Section 6.0 on innovation).

Criterion 2.4: Changes to National Policies and Programs

UNICEF has had considerable success in assisting national and local governments to both develop and to implement policies and programs aimed improving the lives of children and adolescents (74 percent of evaluations reporting satisfactory or highly satisfactory results). Most often this is

the result of ongoing application of technical support and advocacy to both policy development and its implementation. Contributions to changing national policies and programs for children have been noted across all five goal areas of the UNICEF Strategic Plan 2018-2021.

Table 6: Supporting Changes to National Policies and Programs

Goal Area	Selected Examples
Survive and thrive	<ul style="list-style-type: none"> • Development of a national nutrition response task force (Vietnam); • Technical support on the use of Rapid SMS to assist district level health authorities to reach performance targets (Rwanda).
Learn	<ul style="list-style-type: none"> • Technical support to policy development, curriculum development and teacher training in Early Childhood Development (Uganda). • The model implemented in the Initiative has influenced policies and secondary legislation in education. The SAI model is included in the projects promoted by the Ministry of Education (Romania).
Protection from violence and exploitation	<ul style="list-style-type: none"> • Development of national guidelines child protection policies, codes of conduct, minimum standards, and key messages for children and caregivers (Lebanon). • Improvements in legislation for child protection (Kazakhstan).
Safe and clean environment	<ul style="list-style-type: none"> • Support to development of policies on health, nutrition and water and sanitation in humanitarian contexts (Sudan).
Equitable chance in life	<ul style="list-style-type: none"> • The project was instrumental for the Ministry of Youth and Sports ddeveloping a national youth policy and corresponding programs to rehabilitate young people and successfully adopted a national life skills and employability program (Egypt).

There is no single factor which stands out as key to ensuring success in advocacy and technical support for positive change in national policies and programs aimed at achieving the SDGs for children and adolescents. Most evaluation reports noted that UNICEF was able to bring together a number of critically important approaches to ensure it can exert a positive influence on national policies and programs. These include:

1. Consistent advocacy focused on specific national issues such as the need to address teenage pregnancy (Thailand);
2. Credible technical capacity to assist in policy development, often based on experience in support of program implementation (DPR Korea);
3. Inter-sectoral consultations with key stakeholders at national and sub-national level (Bosnia and Herzegovina);
4. Strong and enduring partnerships with national ministries, UN specialized agencies, national research institutions and active international and national NGOs (Paraguay, Ghana).

A theme of the evaluation reports reviewed is **the necessity for long-term, consistent and persistent advocacy to achieve important changes in national policy for children and**

adolescents. While progress may be slow, as in advocating for changes in government policy on justice system engagement with children, it is not negligible when measure over time.

4.3 Cross-Cutting Themes

Summary

UNICEF evaluations have improved over time in consistently and substantively addressing program effectiveness in addressing gender equity (GE). More importantly, UNICEF supported programs are more effective in addressing Gender Equity as assessed by the 2020 DER when compared to earlier reviews. This reflects greater attention to Gender Equity under the GAP 2018 – 2020 and is consistent with the stated objectives of the current Strategic Plan. The recent GAP evaluation credited the two plans with establishing the needed architecture for effective support to Gender Equity at UNICEF and for increasing the profile and presence of Gender Equity initiatives in UNICEF supported programs. It also pointed to the need for greater ambition in UNICEF support to Gender Equity for children and women.

The body of evaluations reviewed in 2020 also support the need for continued improvement to build on the 58 percent of evaluations reporting positive results for effective support to Gender Equity: mainly by ensuring that program development and design processes include significant, systematic and explicit efforts to identify and address barriers to Gender Equity. Nonetheless, improving results in effective support to Gender Equity provide important evidence that UNICEF is able to shift resources and attention to better address strategic priorities over time.

Overall results for promoting equity among marginalized and vulnerable children and adolescents are largely positive for the 2020 review. Quite often these positive results are found for programs and activities targeting inequities experience by children with disabilities or adolescent girls. Where UNICEF supported programs fail to effectively address inequities, this often reflects a lack of attention to the needs of vulnerable and marginalized children and adolescents during program design and/or a lack of participation by targeted communities during the design and implementation phases.

Coverage: Cross Cutting Themes

The pattern of coverage for cross cutting themes is the same for the 2020 review as in 2016. In each case, coverage for sub-criteria 3.1 (Gender Equity) and 3.3 (addressing issues of equity) was rated strong with valid findings reported in fifty or more evaluations. Unfortunately, the pattern of very low coverage for sub-criterion 3.2 on environmental sustainability remains very weak (as in 2012 and 2016) with just five evaluation reports reporting findings. As programming which recognizes and accounts for climate change is called for in the current Strategic Plan, it is likely that coverage of environmental sustainability will be covered more often in evaluations commissioned after 2018.

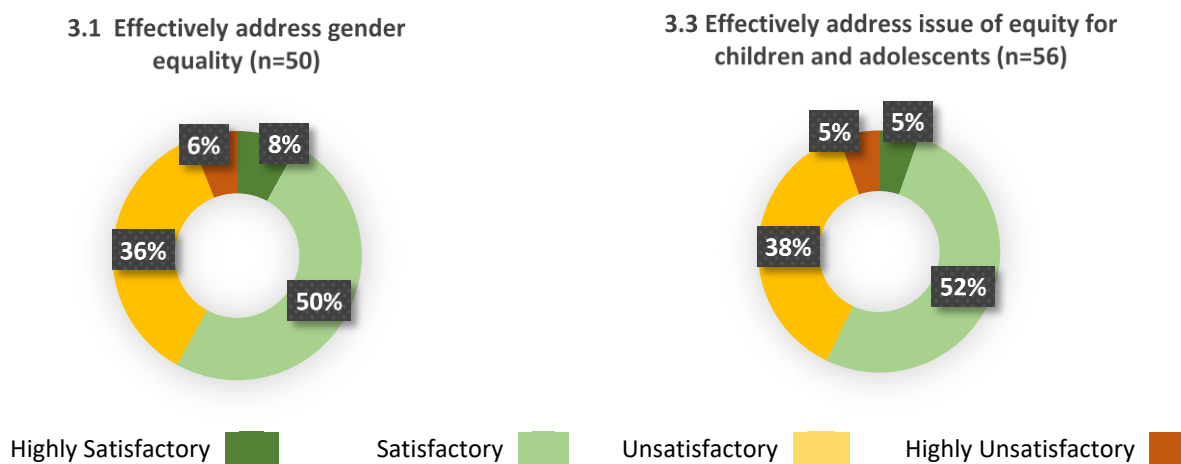
Key Findings: Cross Cutting Themes

Criterion 3.1: Gender Equity

Each of the criteria for cross-cutting themes presents a different pattern of results and findings. The trend in both coverage and results for effectively addressing Gender Equity is both consistent and positive across the 2012, 2016 and 2020 reviews. In 2012, coverage of Gender Equity in UNICEF evaluation reports was only moderate and results were disappointing with only 47 percent reporting positive findings. By 2016, coverage of Gender Equity was rated strong but results were still disappointing with only 48 percent of evaluations finding UNICEF performance on Gender Equity either satisfactory or highly satisfactory. This has changed again in the latest review (2020). With strong coverage of how UNICEF addresses Gender Equity, 58 percent of evaluations reported positive findings with eight percent indicating UNICEF performance on GE was highly satisfactory. This pattern is even more significant given that the analysts reviewing each evaluation report indicate that coverage of Gender Equity was more consistent, more in-depth and more rigorous in evaluations reviewed in 2020 than those sampled for the two previous reviews.

. UNICEF efforts to improve effectiveness in supporting efforts to address Gender Equity are a reflection of its importance as an organizational priority. The original UNICEF Gender Action Plan 2014-2017 was given fresh impetus by the results of the 2016 Effectiveness Review which noted increased attention to Gender Equity but not yet improved results. The original action plan was replaced by the UNICEF Gender Action Plan 2018 – 2021 with established targets in GE for each Goal Area of the Strategic Plan.

Figure 3: Results for Cross Cutting Themes



UNICEF supported programs and activities contributed to advancing Gender Equity across a broad spectrum of interventions contributing to all five UNICEF strategic Goal Areas. This is to be expected given the cross-cutting nature of GE as a results area.

Table 7: Contributing to Gender Equity in UNICEF Goal Areas

Goal Area	Selected Examples
Survive and thrive	<ul style="list-style-type: none"> Addressing discrimination and harassment of female students (Benin).
Learn	<ul style="list-style-type: none"> School attendance initiatives which were designed and implemented taking account of the possible effects of gender differences on attendance (Romania). Life Skills Education (LSE) integrating issues relating to Gender Equity (India). Developing and promoting guidelines and standards for curriculum to adequately address Gender Equity (Eswatini).
Protection from violence and exploitation	<ul style="list-style-type: none"> Mobile approaches to providing GBV and child-protection services in hard to reach areas in humanitarian contexts (Lebanon) Building capacity for training institutions to design, run, monitor and evaluate training in child protection which is gender-sensitive and rights-based (Malawi).
Safe and clean environment	<ul style="list-style-type: none"> Ensuring secure, sanitary and gender-appropriate sanitation facilities and access to quality water for drinking, cooking and maintaining personal hygiene (Sudan).
Equitable change in life	<ul style="list-style-type: none"> Addressing gaps in the incorporation of gender-specific elements in EPI activities to ensure access to women care givers and their children (Lebanon).

The recently completed UNICEF evaluation of the Gender Action Plans identified similar areas of positive results in GE¹² including: for example, support to girls’ secondary education; addressing gender-based violence in emergencies; addressing women’s roles in water, sanitation and hygiene; and ensuring that national cash transfer programs address inequities in women’s incomes. On the other hand, the GAP evaluation also pointed to the need for greater ambition by UNICEF in supporting GE: “Both GAPs were constrained by limited aspiration and by insufficient accountabilities, which in turn undermined broad-based ownership. Their design spoke to the

Box 11: Gender Equality and Child Protection in Malawi

The NCPS had pushed the country’s agenda on gender equality forward. The NCPS mainstreamed gender and rights issues into the overall strategy including through statements such as, “building the capacity of training institutions to design, run, monitor and evaluate competency-based training that is both gender-sensitive and rights-based” A positive aspect of the NCPS was the fact that it included attention to violence against women and not just against children. This was important because when violence against women occurs, children could also be affected emotionally and physically. The findings also revealed that the NCPS had contributed to some socially transformative changes, particularly in the sphere of encouraging girls to stay in school and attention to combatting child marriage, while much work remains to be done.

Source: Evaluation of Malawi Child Protection Strategy 2012-2018, p.60

¹² UNICEF, Realizing Potential: Evaluation of UNICEF’s Gender Action Plans, UNICEF Evaluation Office, December 2019. p, 59,60.

UNICEF operating model, but they lacked sufficient drive or ambition to fully permeate UNICEF’s core programming areas.”¹³

Criterion 3.2 Environmental Sustainability

As already noted, UNICEF evaluations have not assessed the environmental sustainability of supported programs and activities in any significant numbers. It is also worth noting that very few evaluation reports make any reference to climate change (seven of the 62 evaluations in the sample in 2020). Where evaluations do address climate change, they do so by noting that climate change represents an important contextual factor rather than a challenge to be directly addressed by UNICEF supported programs. However, there were a few references to programs aimed at addressing some of the consequences of climate change. In Vietnam, for example, UNICEF supported efforts to improve access to piped water and larger water storage containers for rainwater as a response to expected adverse changes in climate. In contrast, in Malawi, a UNICEF supported latrine program was criticized for making use of local material and logs which contributed to deforestation. In general, however, the issue regarding the environmental effects of UNICEF programming is a lack of attention paid by most evaluation reports.

Criterion 3.3 Effectively Addressing Issues of Equity

As already noted, UNICEF has seen improvement in both coverage and results for effectively addressing gender quality which is one of the key components of addressing equity for children and adolescents. However, equity issues go beyond questions of Gender Equity into other dimensions of marginalization and discrimination based on extreme poverty, minority status, rural isolation, children with disabilities and the general challenge of reaching marginalized adolescents. Sub-criterion 3.3 was introduced into the 2016 review and continued in 2020 in an effort to track UNICEF effectiveness in addressing the needs of children and adolescents suffering from these forms of marginalization.

In 2020 UNICEF results for this criterion were positive with 57 percent of evaluations finding UNICEF had achieved satisfactory or highly satisfactory results in addressing issues of equity. However, this represents a ten percent decline from the positive results found in the 2016 review. This may have resulted from a more precise definition of equity in the terms of reference of UNICEF evaluations and a higher level of ambition expected when evaluating UNICEF support. At the same time, it is notable that evaluations provide positive examples of efforts by UNICEF to address inequities across all five Goal Areas.

Table 8: Contributions to Addressing Inequities

Goal Area	Selected Examples
Survive and thrive	<ul style="list-style-type: none"> Inclusion of children from different tribes and backgrounds in toy making activities (Uganda)
Learn	<ul style="list-style-type: none"> Inclusion of indigenous people in results targeting and follow up (Argentina) Coverage of student candidates across all caste groups (India)

¹³ UNICEF, Realizing Potential: Evaluation of UNICEF’s Gender Action Plans, UNICEF Evaluation Office, December 2019, p.7.

Goal Area	Selected Examples
	<ul style="list-style-type: none"> Providing extra resources for transportation and enhancing access to teachers for Roma children and employing young people with learning disabilities (Albania, Macedonia and Serbia).
Protection from violence and exploitation	<ul style="list-style-type: none"> Maintaining court and policing system with child protection-based approach (Malawi)
Safe and clean environment	<ul style="list-style-type: none"> Improved social equity by strengthening WASH facilities (India)
Equitable chance in life	<ul style="list-style-type: none"> Effective targeting of Cash Transfer Programs (CTP) to reduce inequities between households (Nigeria). Proactive identification of children at risk including children with disabilities and increased effort to link to available welfare support, such as the Disability Fund (Ghana)

There were two groups which received a relatively higher level of attention in the evaluation reports reviewed in 2020: children with disabilities and adolescents. This seems to be a clear reflection of a greater emphasis by UNICEF on support to improving equity for both groups.

Children with Disabilities

Over half of the evaluations reporting findings on UNICEF support to strengthening equity made a direct reference to efforts to address the needs of children and adolescents with disabilities. Further, 27 evaluation reports made specific reference to program features designed to address inequity for children with disabilities, including examples from all five Goal Areas.

Table 9: Reaching Children with Disabilities

Goal Area	Selected Examples
Survive and thrive	<ul style="list-style-type: none"> Mobile clinics to enhance access to health care for those usually excluded from the system who live beyond the 5-mile radius that surrounds the closest health post (Niger).
Learn	<ul style="list-style-type: none"> Planning schools designed to be accessible for students with disabilities (Palestine). Including children with disabilities in planning to support access to education (Romania).
Protection from violence and exploitation	<ul style="list-style-type: none"> Advocacy for rights of children with disabilities (Thailand).
Safe and clean environment	<ul style="list-style-type: none"> Adaptation of sanitation facilities for persons with disabilities (Lebanon).
Equitable chance in life	<ul style="list-style-type: none"> Program planning documents which clearly identified needs of children with disabilities (Benin). Guidelines for child assessment and creation of individual development plans for children with disabilities (Croatia).

Goal Area	Selected Examples
	<ul style="list-style-type: none"> Capacity building with government agencies to improve effectiveness in working with children and adolescents with disabilities (Nepal).

Adolescents and the Second Decade of Life

Approximately half of the evaluation reports reviewed (44 percent) in 2020 make reference to UNICEF support to policies and programs aimed at reaching adolescents. UNICEF support to addressing the needs of adolescents was most often a mix of advocacy for effective policies and strategies and direct support to programs and activities aimed at adolescents.

Table 10: Contributing to Equity for Adolescents

Goal Area	Selected Examples
Survive and thrive	<ul style="list-style-type: none"> Advocacy for increased awareness of youth at risk for HIV and for a national strategy to address HIV transmission among adolescents (Thailand). Support to Second Chance Centres for adolescent girls who have left school due to pregnancy or early marriage (Nigeria).
Learn	<ul style="list-style-type: none"> Skills development programs aimed to empower 13 to 18-year-old adolescents (Egypt).
Protection from violence and exploitation	<ul style="list-style-type: none"> Efforts to reduce or eliminate child marriage (Guinea Bissau, Somalia, India, Bangladesh, Nigeria). Training on adolescent psychology and working with adolescent prisoners (Kyrgyzstan)
Safe and clean environment	<ul style="list-style-type: none"> No example found
Equitable chance in life	<ul style="list-style-type: none"> Engagement of mentally-disabled and empowering young people and employing young people with mental disability (Thailand, Egypt).

Of course, not all UNICEF supported programs are effective in addressing the needs of adolescents. A few reports noted that programs failed to include adequate measures to target programming to adolescents.

Box 12: Inadequate Attention to Needs of Adolescent Girls in Combatting GBV in South Sudan

Although adolescent girls are a group which is at very high risk of GBV in general and particularly in South Sudan, GBV programming (both specialized and integrated) has not adequately focused on the needs of adolescent girls, including in WFS and Demobilization, Disarmament, and Rehabilitation (DDR) initiatives. CMR trainings currently include only basic information about caring for child and adolescent survivors.

Source: Multi-Country Real Time Evaluation of UNICEF Gender-Based Violence in Emergencies Programs: South Sudan Country Report, 2016. P.8

Changes Over Time

The most important trend for cross-cutting themes over time has been the steady improvement in both coverage and results for sub-criterion 3.1 on the extent UNICEF support effectively addresses Gender Equity. This is consistent with the application of attention and resources over two sequential Gender Action Plans (GAP) at UNICEF: the first from 2014 to 2017 and the second from 2018 to 2021.

Contributing Factors: Cross Cutting Themes

Criterion 3.1: Gender Equity and the UNICEF Gender Action Plan

In 2019, UNICEF published an evaluation of the 2014 to 2017 and 2018 to 2021 GAPs. The GAP evaluation identified the most important steps taken by UNICEF under the GAPs in order to strengthen effectiveness in supporting GE. It concluded that UNICEF had used the Action Plans to: “build some important foundations for UNICEF’s Gender Equity work. They provided a valuable organizational framework for UNICEF’s programmatic efforts on gender and helped build UNICEF’s gender architecture and systems from a limited base. They also succeeded in securing resources for Gender Equity under GAP 1; and in enabling the organization to dedicate efforts and attention to some issue-based (targeted) priorities.¹⁴

At the same time, the GAP evaluation found that UNICEF needs to do more to build organizational accountability for effectiveness in addressing Gender Equity. The evaluation report also stresses the need for UNICEF to update its approach to Gender Equity to better reflect current contexts and to “*significantly upscale its ambition for Gender Equity, commensurate with its status as the world’s defender of child rights.* (p.8)”

Factors Contributing to Positive Results in GE

Most of the evaluation reports reviewed in 2020 identify an explicit effort to identify and address barriers to Gender Equity during program design as the single most important factor contributing to positive results. Inclusion of women in participatory program planning and implementation committees and working groups was also an important positive contributing factor (Cameroon, Guinea Bissau, Kyrgyzstan, Myanmar). Some evaluations also noted that UNICEF programs actively supported involvement of men and boys in efforts to address Gender Equity (Uganda, Paraguay, Lebanon, Croatia). Some evaluations pointed to limited engagement of men as a weakness in addressing Gender Equity. A final, significant negative factor for effectiveness in Gender Equity was the absence of reliable sex-disaggregated data in program monitoring systems (Vietnam, Rwanda, Bosnia and Herzegovina, Palestine, Kazakhstan).

The 2019 GAPs evaluation identified a series of organizational requirements that would need to be addressed to move UNICEF to a higher level of results achievement in GE, including: a revised policy architecture, phased institutional strengthening for GE, embedding GE into country and regional program planning, and ensuring leadership and accountability for ownership.¹⁵

¹⁴ Evaluation of UNICEF’s Gender Action Plans, p.8.

¹⁵ Evaluation of UNICEF’s Gender Action Plans, p.9

Criterion 3.2: Effectively Addressing Issues of Equity

The single most important factor supporting effective UNICEF support to issues of equity is an explicit consideration of the needs of vulnerable and marginalized children and adolescent during program design (cited in 30 percent of evaluation reports). The second most frequently cited positive factor was inclusion of explicit targets for reaching vulnerable children during program implementation, including adolescents and children with disabilities. Finally, adapting innovative program approaches to reaching vulnerable children and adolescents was also noted as an important positive factor in promoting equity. Examples include the use of mobile health teams (Niger, Ethiopia) and cash transfers targeted to vulnerable groups (Nepal, Malawi).

Where UNICEF supported programs fail to address the needs of vulnerable and marginalized children and adolescents, evaluations most often point to inattention during program design as the main limiting factor. This often combines with the absence of effective targeting measures (Sudan, Pakistan, Togo, Jordan) for directing program outputs to the most vulnerable.

4.4 Sustainability

Coverage: Sustainability

All three sub-criteria relating to sustainability were rated strong in the level of coverage provided by the sample of evaluations reviewed in 2020. This is consistent with the 2016 review and an improvement on the moderate level of coverage found in the 2012 review.

Summary

Sustainability, especially financial sustainability, has been a consistent challenge for UNICEF supported programs and activities across all three effectiveness reviews (2012, 2016 and 2020). It remains an important challenge in 2020. In contrast, a clear positive area has been the sustainability of UNICEF efforts to strengthen the capacity of partner institutions and communities with sustained improvement across all three reviews. Most often this reflects careful consideration given to assessing and strengthening the policy development and service delivery capacity of partners during program design. It also reflects a longstanding history of effective partnership by UNICEF. UNICEF has also been able to achieve positive results regarding efforts to improve the enabling environment for development, often by linking programs to existing national institutional and policy structures.

Evaluation reports reviewed in 2020 link the persistent and perhaps worsening problem of ensuring financial sustainability of programs to:

- Lack of planning for budgeting by national governments sometimes combined with planning assumptions regarding sources of sustainable financing which are unrealistic;
- The need for more effective planning for financing program activities as they are taken to scale; and,
- Inadequate or missing exit plans and strategies for the period following UNICEF support.

Evaluations also point to persistent problems in securing financing for, in particular, sustained health systems strengthening at national level – either from national budgets or from multilateral and bilateral sources of financing. In 2018 UNICEF published an evaluation of the UNICEF contribution to Health Systems Strengthening. The evaluation reported that: “UNICEF needs to be more systematic and follow a clear, step-wise approach to replication of global models and scaling interventions that have been proven effective” as part of an overall effort to make support to health system strengthening more sustainable.¹⁶

Key Findings: Sustainability

Ensuring the sustainability of results achieved through support to programs and activities is an enduring challenge for UNICEF. In fact, sub-criterion 4.1 (the likelihood that program benefits would continue after completion) saw the portion of positive findings diminish between the 2012 review (51 percent) and the 2016 review (36 percent). It fell further in the three years covered by the 2020 review when only 32 percent of sampled evaluations reported satisfactory or highly satisfactory results for sub-criterion 4.1. As noted below, the most common factor in reduced sustainability has been ongoing difficulty in securing medium to longer-term program funding by national governments **in the period following UNICEF support.**

Box 13: Capacity Development for Sustainable Child Protection in Kyrgyzstan

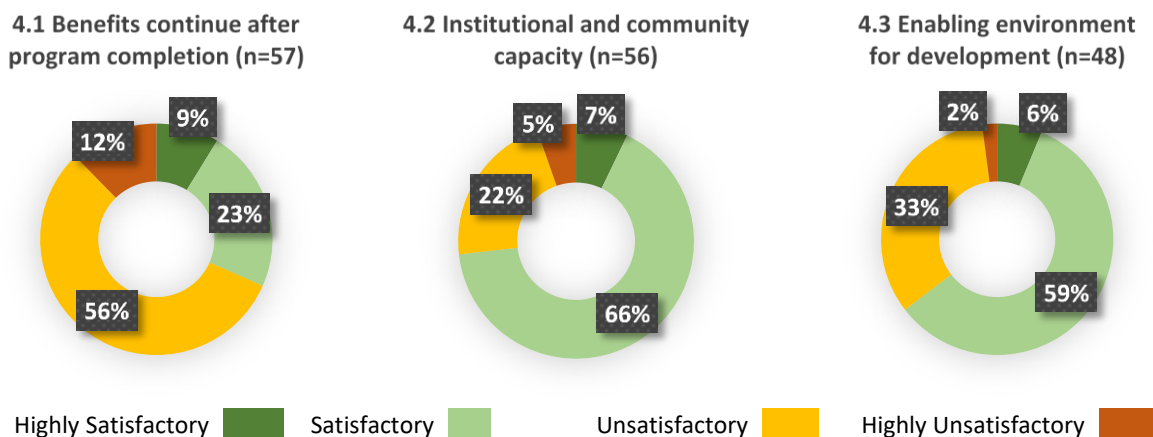
Hundreds of judges, prosecutors, police officers, social workers, correctional workers and lawyers were trained during the Program. The training of judges had a significant impact on justice for children. Every trial court now has at least one judge trained in justice for children. The training of police officers also has had a significant positive impact on the way juvenile offenders and child victims are treated by inspectors and investigators.

Source: Summative evaluation of implementation of the State Program on Justice for Children in Kyrgyzstan 2014-2018, 2019 (p7)

However, there is an aspect of sustainability where evaluations find improving results over time. The 2020 review found that 73 percent of evaluations reported that UNICEF supported programs were sustainable **from an institutional and community capacity perspective.** This represents steady and significant improvement from the 2012 (48 percent) and 2016 (64 percent) reviews. This improvement suggests consistent and increasing attention by UNICEF to the need for capacity development action as a basis for effective programming.

¹⁶ UNICEF Evaluation Office, Evaluation of UNICEF Contribution to Health System Strengthening, 2019, p.50.

Figure 4: Results for Sustainability



UNICEF also continues to be reasonably successful in supporting efforts to strengthen national policies and programs and, as a result, to strengthen the enabling environment for development at a national level (sub-criterion 4.3). The frequency of positive findings for this criterion declined slightly to 64 percent in 2020 from 74 percent in 2016. This may also be a result of a stronger focus on the operational aspects of UNICEF programs in recent evaluations rather than a focus on advocacy and policy engagement.

Changes Over Time

The problem of achieving sustainable benefits for target group members through UNICEF supported programs has remained an important challenge across all three effectiveness reviews. In fact, the weighted average score for sub-criterion 4.1 declined from 2012 to 2016 and remained very low in 2020. This may reflect increasing difficulty on the part of national governments to secure public finances to sustain program operations when UNICEF support is ended. In contrast, findings for the sustainability of institutional and community capacity supported by UNICEF improved steadily over the three reviews. This improvement could be grounded in a continuing focus on strengthening public service delivery systems (especially including health systems) on the part of UNICEF and other development partners.

Contributing Factors: Sustainability

Sub-criterion 4.1: Sustainability of Program Benefits

The most commonly reported positive factor supporting the sustainability of program benefits into the future, cited in just over half of the evaluations with positive findings, concerns efforts by UNICEF to ensure that programs were mainstreamed into national strategies and budgets and/or linked to community actions at the sub-national level. Examples of this positive approach include:

- Integrating emergency GBV efforts into ongoing development programs (South Sudan);

- Embedding action to improve children’s access to equitable justice practices in national legislation (Kazakhstan); and,
- Embedding a nursery education program into the sector plan for education with support from the national budget (Guyana).

Another important factor for improved sustainability is noted whenever UNICEF supported programs are able to include explicit plans for replication and scale-up during the pilot phase. Evaluations carried out in India, Nigeria, Egypt and Malawi all provided examples of how scalable programs were built on plans developed during smaller scale pilots.

Almost all evaluation reports that reported negative findings on sustainability pointed to the absence of a source for adequate ongoing funding as the key limiting factor. Sometimes, as in Algeria, this reflects the lack of a plan or effort for mobilizing resources after UNICEF support is withdrawn. In other cases, planned financial support from community members was not forthcoming and was deemed unrealistic by the evaluators (Benin, Cameroon). It is important to note that, while the envelope of financial resources available to UNICEF has grown during the period under review, this sub-criterion refers to the likelihood of sustainability in the period after UNICEF support has been withdrawn.

There are other, more specific, examples of factors that undermine the sustainability of program benefits after UNICEF support is ended. These include:

- **Absent or inadequate exit strategies** for ending UNICEF support (Jordan, Palestine, Thailand, Lebanon, Ethiopia);
- **Unrealistic assumptions on financial support** from government agencies, community members or the private sector (Zimbabwe, India, Ghana, Nepal);
- **Staff turnover** and failure to adequately address **staff capacity** (Romania, Ethiopia); and,
- **Inadequate stakeholder involvement in program design** and implementation (Honduras, Indonesia, Congo).

Taken as a group, these factors suggest that program planning processes did not take adequate account of the need for exit strategies grounded in stakeholder ownership and based on realistic assumptions about continuing funding.

Sub-Criterion 4.2 Institutional and Community Capacity

Forty percent of evaluations with positive results for sustaining institutional and community capacity point to UNICEF engagement in ensuring adequate planning for training, technical assistance and other aspects of capacity development as a critical factor. Examples of targeted capacity development actions include:

- Technical guidance and capacity development assistance to practitioners in ECD (Uganda);
- Training of volunteers and volunteer coordinators (Honduras);
- Capacity building among Community Based Organization (CBO) and MoH staff (Jordan);

- Innovative attention to capacity building for child protection with staff of security organizations (South Sudan); and,
- Establishing and developing capacity for community-based structures to prevent malnutrition and support vaccination programming (Mali).

Other frequently cited positive factors include:

- Integration of programming structures into sub-national administrative bodies (Korea);
- Long-standing UNICEF leadership and presence (Nepal);
- UNICEF commitment to support and strengthen existing structures for programming (Zimbabwe, Ethiopia); and,
- UNICEF support and engagement with partnerships, coalitions and alliances to strengthen and sustain program capacity (Thailand, Bangladesh, Malawi, Palestine).

Sub-Criterion 4.3 Contributing to the Enabling Environment

The single most important factor supporting the sustainability of UNICEF efforts to influence the enabling environment for development, cited in 11 evaluation reports (27 percent) is a direct link from UNICEF support to a clear institutional and policy framework supporting positive change for children. This in turn, most often builds on a credible history of UNICEF support and strong national and local partnerships. Examples include:

- A strong institutional and policy framework for scaling UNICEF support (Malawi);
- A strong working relationship and a history of supporting policy reform (South Sudan); and,
- Embedding UNICEF supported education programming within the National Education Plan (Guyana).

The second most frequently noted positive factor supporting an improved enabling environment for development that benefits children and adolescents (9 evaluation reports, 22 percent) can be found in UNICEF efforts to promote national ownership and local political support and commitment. The quality and durability of national and local ownership was specifically referred to in most evaluation reports.

4.5 Efficiency

Summary

Efficiency and value for money issues are becoming more important and are receiving more attention in UNICEF evaluations. It is true that the raw numbers on coverage for efficiency criteria have not notably increased from 2012 and 2016 to the 2020 effectiveness reviews (see table 1). However, efficiency issues are covered. It is also important to acknowledge that UNICEF staff and their partners can and do plan and implement programs from a strong value-for-money

perspective. When this happens, UNICEF is able to support programs with relatively low unit costs and significant reach. at greater length and with more in-depth analysis in evaluations published in recent years.

However, greater attention to efficiency issues in evaluations has not been accompanied by more positive results. Cost efficiency was found satisfactory or highly satisfactory in only 47 percent of the evaluations reviewed in 2020, compared with 64 percent in 2012. UNICEF also faces persistent challenges in implementing programs and achieving objectives in a timely manner. Persistent challenges include:

- Over-ambitious and unrealistic timelines
- Administrative delays in approval, contracting and start-up
- Delays in disbursements to implementing partners, and
- Supply chain management problems.

For some UNICEF supported programs and activities, efficiency problems originate with weak program planning which results in unrealistic projections of required financial, human and system resources (and inadequate timelines).

Coverage: Efficiency

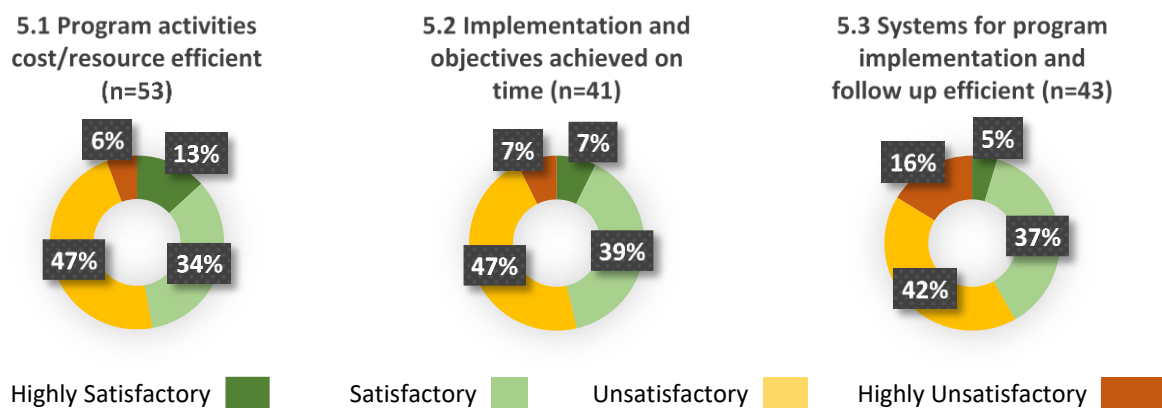
Coverage of sub-criterion 5.1 on the cost efficiency of UNICEF supported interventions was rated strong for the 2020 review with 53 of the 62 reviewed evaluations presenting credible findings. Sub-criteria 5.2 and 5.3 (timeliness of delivery and efficiency of programming systems) were covered to the moderate level in 2020. This pattern of strong coverage of sub-criterion 5.1 and moderate coverage for 5.2 and 5.3 has remained consistent across all three effectiveness reviews. This pattern indicates that efficiency criteria are seldom central to the focus of the evaluations reviewed. As a result, when efficiency is addressed in evaluation reports, it most often defined in narrow terms of ensuring use of low-cost inputs rather than the efficiency and timeliness of programming systems.

Key Findings: Efficiency

The cost and resource efficiency of UNICEF supported programs and activities remains a difficult challenge as assessed by evaluations reviewed in 2020. Less than half (47 percent) of the relevant evaluations found that programs are efficient in terms of costs and resources used. This represents a decline from the 62 percent figure achieved in 2016, which, in-turn, may reflect rising expectations regarding cost-efficiency on the part of evaluations carried out from 2016 to 2019.

A similar pattern is evident for the remaining sub-criteria on efficiency. Only 46 percent of evaluations reported that UNICEF supported programs were implemented and objectives achieved on time (sub-criterion 5.2). Finally, only 43 percent of evaluations found that UNICEF systems for implementation and follow-up (including systems for engaging staff, procuring inputs, arranging logistics and securing payments) were efficient (Sub-criterion 5.3).

Figure 5: Results for Efficiency



It seems credible that relatively negative findings for sub-criteria 5.3 on the efficiency of systems for program implementation would, in turn, effect findings for sub-criterion 5.2 on timely programming. If systems for hiring contracted staff, procuring program inputs and disbursing funds to implementing partners are inefficient this could result in delays in program implementation.

There are, however, examples of UNICEF providing support to programs and activities found to be cost efficient in just under half the evaluations (49 percent) reviewed. Examples include:

- Efficient use of human resources in UNICEF supported programs compared to similar initiatives in the same country (Lebanon, Guinea Bissau, Honduras);
- Strict adherence to realistic program budgets (Egypt, Bosnia and Herzegovina, Malawi, Togo, Benin, Palestine; and,
- A focus on quality, low-unit cost interventions as a means of increasing scale and reaching more target group members (Honduras, Jordan, Ethiopia).

Box 14: Cost Effective Psycho-Affective Care Services in Honduras

The Return of Joy offered a psycho-affective care service to a massive number of vulnerable children and adolescents at low cost, which proved to be efficient. The Return of Joy methodology proposes that the strategy be implemented by adolescent volunteers. In this sense, the participation of adolescent volunteers was very efficient, as more than 500 adolescents offered their time free of charge for the play sessions.

Source: Evaluation of the implementation of the Return of Joy strategy for migrant children in Honduras, 2017. (p. 40)

Aspects of UNICEF programming systems found wanting in the reviewed evaluations include:

- Insufficient resourcing in terms of allocated staffing, allotted time frames and funds compounded by a lack of standardized procedures for program management (Thailand);
- Weak financial planning and delayed contracting (Burkina Faso);
- Poor risk planning and risk management (Mali);
- Delayed financing and disbursements (Mali, India, Bangladesh, Indonesia, Nigeria);
- Problems in procurement and management of supply chains (Sudan, Pakistan, Vietnam, Ethiopia and Palestine); and,
- Delays and inadequacies in hiring and fielding staff and managing human resources (Togo, Nepal, Guinea Bissau, Niger, Malawi, Vietnam).

Changes Over Time

The analysts conducting the 2020 review (who also reviewed evaluation reports in 2016) report that the current sample of evaluations published between January 2016 and December 2019 provide more in-depth coverage and more systematic consideration of efficiency questions, particularly questions of unit costs when compared to reports sampled for earlier reviews.

However, while the quality of analysis for cost effectiveness criteria improved, the number of evaluations reporting positive findings declined across all three efficiency sub-criteria between 2016 and 2020 (Table 7). This may be a result of closer scrutiny against more stringent indicators in recent evaluations.

Contributing Factors: Efficiency

Sub-criterion 5.1 Cost and Resource Efficiency

Many evaluation reports with positive findings (15 cases, or 60 percent) emphasized that UNICEF supported programs were cost efficient because UNICEF and its partners were highly cost conscious and took a “value for money” perspective when designing and implementing interventions. In Zimbabwe for instance, program managers reduced training costs by using localized personnel for training and renegotiating fees.

Evaluation reports with positive findings for cost efficiency, often (9 cases, 37 percent) point to the use of cost-benefit analysis to maximize benefits in relation to costs. Other factors promoting efficiency include:

- A staffing and training approach which provided UNICEF supported programs with access to low cost human resources for implementation, sometimes through use of volunteers and community contributed labour (Lebanon, Guinea Bissau, South Sudan);
- A focus during program planning on identifying program inputs with low unit costs (Honduras, Pakistan); and,
- Establishing realistic program budgets from the beginning of the planning and approval process and holding managers accountable for adherence to these realistic budget projections (Togo, Benin, Palestine, Congo). It is important to note that where budgets are

not realistic and do not cover the minimum requirement for effective implementation it soon becomes impossible to hold managers accountable for remaining “on budget”.

The most common negative factor leading to reduced cost efficiency has been **lack of credible information on costs and expenses**. Without access to cost data which can be compared to budgets and projections, UNICEF and its partners have little ability to identify excess costs and improve efficiency. Other significant negative factors for cost effectiveness include:

- Overlaps and duplication in service provision resulting from a lack of coordination among implementing partners (Palestine, Algeria, Jordan, Malawi); and,
- Use of inefficient mechanisms to deliver services or support program activities such as the use of commercial banks to distribute cash to beneficiaries in Kenya, resulting in high transport costs for target group members.

Sub-Criterion 5.2 Implementing Programs and Achieving Objectives on Time

Evaluations with positive findings on the timeliness of UNICEF supported programs and activities identified three significant factors which helped to ensure on-time start-up and delivery of services:

- Use of pre-existing partner networks to eliminate or mitigate delays (Vietnam, Zimbabwe)
- Ensuring program implementation systems are flexible enough to address new or emerging priorities without introducing delays (Albania) and,
- Using a single mechanism to administer and fund support to the health sector in order to avoid separate, vertical programs supported by different external partners as for the Health Transition Fund in Zimbabwe.

On the other hand, evaluations reporting negative findings for on-time program start-up and service delivery pointed to delays in disbursing funds to implementing partners and the single most important contributing factor. Other factors causing delays include:

- Poor planning with lack of attention to the implementation timeline (Kazakhstan, Kenya, Malawi). As with unrealistic budgets, overly-ambitious program timelines under-mine efforts to support good program management and hold managers to account;
- Delays in approval and contracting (Jordan, Togo, Nepal) and,
- Weaknesses in supply chains and supply chain management (Ethiopia, Sudan, Pakistan).

Sub-criterion 5.3 Efficient Systems and Procedures for Program Implementation

Where evaluation reports find UNICEF program approval, implementation and follow up systems and processes are efficient they often (6 cases) point to the strength of coordinating mechanisms and ongoing partnerships. Another important factor (7 cases) in efficient programming has been UNICEF's willingness to show leadership and to encourage UNICEF staff to help overcome bottlenecks and demonstrate commitment (Argentina, Benin, Indonesia).

4.6 Using Evaluation and Monitoring to Improve Effectiveness

Summary

Starting from a solid base in the 2012 review, systems and processes for evaluation have been improving over time with 77 percent of evaluations reported positive findings on the effectiveness of evaluation at UNICEF. This finding is consistent with the improvement in evaluation report quality scores achieved over the 2012, 2016 and 2020 effectiveness reviews. UNICEF has also continued to strengthen the use of evaluations to improve development and humanitarian effectiveness as measured through the number and quality of detailed management responses. However, monitoring and reporting of results has been a persistent weakness with evaluations reporting positive findings in just around 40 percent of cases in all three reviews. Not surprisingly, given findings on results monitoring, the promise of Results-Based Management as a systematic approach to improving effectiveness has not been realized at UNICEF as reported in evaluations from the last two review periods (2016 and 2019).

A persistent problem in both results monitoring generally and in RBM has been lack of attention to defining and monitoring results at all levels including activities and outputs but, especially in terms of outcomes. This further relates to the absence of a credible program theory of change developed during the program design phase. Another problem for RBM has been lack of attention or inability to establish an adequate baseline for assessing results (and subsequently tracking relevant indicators to compare with established targets).

It is important to note that UNICEF has made a considerable effort over time to strengthen systems for results monitoring and reporting up to and including the reorganization of the organizational structures and accountabilities in 2019. The development of the UNICEF Strategic Plan 2018 – 2021 also saw the implementation of a comprehensive Results Framework with targets and indicators specified for each Goal and Results Area down to the level of Outputs. These efforts have not yet resulted in strengthened monitoring and RBM systems at the level of individual country programs as reported in evaluations reviewed in 2020.

Coverage: Using Evaluation and Monitoring

Coverage of sub-criteria for evaluation and monitoring remained exceptionally strong for the 2020 review. In every case, sub-criteria were the subject of valid findings in over 50 of the 62 evaluations in the review sample.

Key Findings: Using Evaluation and Monitoring

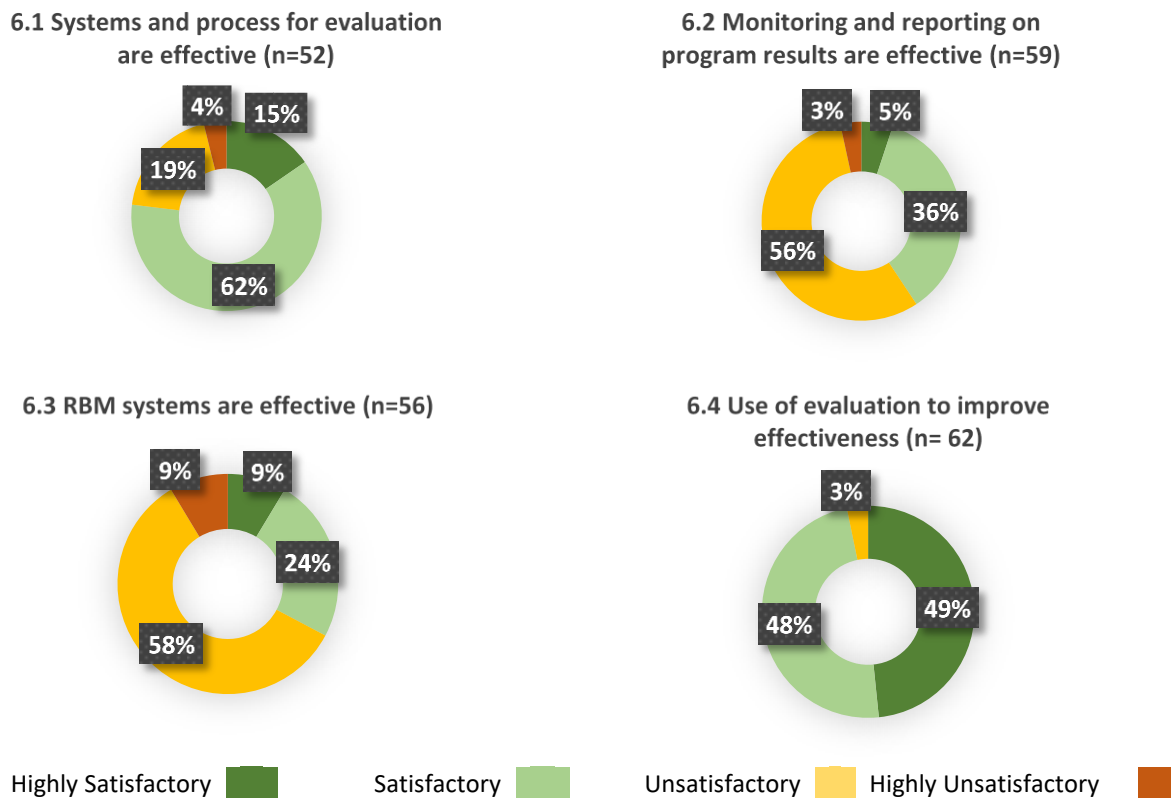
For the overall criterion of using evaluation and monitoring to improve development effectiveness it is best to consider sub-criteria for monitoring together (6.2 and 6.3) before turning to those for evaluation (6.1 and 6.4).

Sub-Criteria 6.2 Effective results Monitoring and Reporting and 6.3 Effective RBM Systems

Despite some positive examples (as noted in Box 14) effective monitoring and reporting and the effective use of results-based management (RBM) systems remained a critical challenge for UNICEF according to evaluations published from 2016 to 2019. For sub-criterion 6.2 on the

effectiveness of monitoring and results reporting systems, only 41 percent of evaluations reviewed in 2020 reported satisfactory or highly satisfactory findings. This was only a marginal improvement over the 38 percent positive result for the same sub-criterion in 2016. Finally, for sub-criterion 6.3 on RBM systems reported results are even more negative. Only 33 percent of evaluations reviewed in 2020 reported satisfactory or highly satisfactory results for the effectiveness of RBM systems. The pattern of poor results for the strength of RBM systems at UNICEF has been consistent across the 2012, 2016 and 2020 reviews.

Figure 6: Results for Evaluation and Monitoring



As with efficiency, it is important to note that negative evaluation findings for one sub-criterion can provide an important explanation for negative findings in another. This is the case for sub-criterion 6.2 on the effectiveness of monitoring and reporting on program results – the 58 percent of evaluation reports with unsatisfactory or highly unsatisfactory results for this sub-criterion provide an important explanation for negative results on sub-criterion 6.3 on effective RBM. If monitoring systems are weak it is evident that RBM systems will also be ineffective.

Box 15: Weak Results Statements Undermine Monitoring in Girls Education

“Results statements for girls’ education were often absent or ill-defined, and there is insufficient evidence from either the desk review or the case studies to determine the extent to which anticipated results were achieved. The lack of adequate results statements and sufficient evidence meant that the evaluation was unable to assess UNICEF’s contribution to attaining gender parity in key education measures, including the MDGs 2 and 3.

Source: UNICEF Evaluation Office, Evaluation of UNICEF Girls Education Portfolio (2009-2015), 2016, p. xi.

Sometimes, weaknesses in monitoring and reporting which feed into poor RBM findings are directly linked to inadequate or vague program results frameworks and definitions (Box 15).

Of course, there are examples of effective monitoring of UNICEF supported programs in the evaluation reports reviewed in 2020. Effective monitoring can and does contribute to learning and corrective actions as expected when monitoring systems and RBM practices are in place (Box 16).

Box 16: Robust and Effective Monitoring of Humanitarian Action in Vietnam

The robust monitoring system established and operationalized by UNICEF went beyond accountability responsibilities and allowed for continuous learning and timely corrective actions through periodic dialogue with partners and national and sub-national levels. (98 monitoring visits were carried out during the project cycle including joint UN and GOV visits). Following UNICEF monitoring visits reports were discussed with national partners and remedial action taken.

Source: External After-Action Review of the Government of Viet Nam and UNICEF emergency response supporting children and women in 10 provinces in Viet Nam affected by drought and salt-water intrusion crisis (2016-2017), 2018 (p35-36)

Efforts to strengthen program monitoring

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

In 2014 a fourth Deputy Executive Director (DED) post was approved and established at UNICEF HQ to oversee work on monitoring field results. A Field Results Group (FRG) was created reporting to the new DED. The FRG assumed the performance management systems and transparency initiatives as well as the program policies and guidance functions formerly the responsibility of the Division of Policy and Strategy (DPS). The DPS was then re-constituted as the Division of Data Research and Policy (DRP) with responsibility for data and data analytics; the strategic plan and corporate reporting; knowledge and civil society partnerships. In 2018 the Repositioning and Realignment Review of the Planning Monitoring and Reporting (PMR) Functions found that these changes had “exacerbated the misalignments in planning, monitoring and reporting with fragmentation of PMR across three Deputy Executive Directors, reducing efficiency and effectiveness.”¹⁷

In early 2019, the Executive director responded to the realignment review with a Decision Memo which created one Program Group under the DED Programs with responsibility for the PMR function at UNICEF. The new Program Group has just three divisions: EMOPA, Programs and the Division of Analysis, Planning and Monitoring (DAPM). The formation of the three program groups was intended to:

- Address the fragmentation and create unity of purpose within the Program group;
- Ensure a common vision, direction and oversight of the PMR functions;
- Strengthen accountability and speed up decision making on PMR;
- Increase the potential for synergies among different functions;
- Establish a clear line of sight between planning and reporting; and,
- Improve the management of global results and formulation of global commitments.¹⁸

Interviews with staff of the DAPM highlight some of the challenges (and progress made) in realigning the PMR system at UNICEF. These include:

- Continuing to refine and improve reporting on results in the Strategic Plan through the system of annual reporting by country offices in response to the Strategic Monitoring Questions;
- Adopting to the increasing focus on climate change in the PMR;
- Tracking and refining UNICEF related SDG indicators;
- Improving data quality and strengthening monitoring and reporting by partners;
- Continuing the evolution and strengthening of the Results Assessment Module (RAM); and

¹⁷ UNICEF, Office Management Plan, July 2019-2021, Division of Analysis, Planning and Monitoring (DAPM). 2019, p.5

¹⁸ Ibid. p.6

- Developing common approaches and cooperating with other UN entities to eventually develop common norms and standards on performance monitoring and reporting.

Finally, it is important to note that the UNICEF Strategic Plan 2018 – 2021 was accompanied by a detailed Theory of Change and a comprehensive Results Framework with results and targets established for each Goal and Results Area and for Outcomes and Outputs with accompanying indicators at all levels. This system was used to provide data for the Mid-Term Review (MTR) of the Strategic Plan published in 2020. It would be reasonable to expect evaluations commissioned after the implementation of the new Results Framework to report different findings on the strength of program monitoring.

Sub-Criteria 6.1 Effective Evaluation Systems and 6.4 Using Evaluation to Improve Effectiveness

The two sub-criteria relating to the effectiveness of evaluation systems and the use of evaluation to improve effectiveness have both seen steady improvement from a solid base over time. For the current review, 77 percent of evaluation reports found UNICEF evaluations systems effective to a satisfactory or highly satisfactory level. This was an improvement on the 70 percent rating in 2016 which, in turn, improved on the 64 percent finding in 2012. Similarly, by 2020, 96 percent of evaluation reports reviewed were accompanied by a management response which included assigned responsibilities and a time-bound plan of action to implement evaluation recommendations.

Changes Over Time

It is important to point out that evaluations published by UNICEF continue to improve in quality over time, providing an important triangulation of findings on improving evaluation effectiveness. Table 8 provides an overview of the distribution of quality scores for the evaluation samples reviewed in 2020, 2016 and 2012. These scores were arrived at using quality scoring criteria and weightings described in detail in Annex E. The same criteria and weights were applied in all three reviews so the quality scores are comparable over time.

The most important trend to note is the increasing number and proportion of evaluation reports clustered in the higher score categories over time. In 2012, only 22.7 percent of evaluation reports received scores in the 36-40 range (with 40 as the maximum) but by 2020 this range accounted for 52.4 percent of all evaluations. The minimum score to be included in the review is 25 points so the lowest possible range is from 25 to 30 points which held 22 percent of all evaluation reports in 2012 but only 16 percent in 2020.

Table 11: Distribution of Evaluation Quality Scores 2020, 2016 and 2012

Score Distribution Categories	2020		2016		2012	
	No. of evaluations	% of evaluations	No. of evaluations	% of evaluations	No. of evaluations	% of evaluations
Max points = 40 (Min required = 25)						
36 – 40	33	52.4%	18	24.7%	15	22.7%

Score Distribution Categories	2020		2016		2012	
	No. of evaluations	% of evaluations	No. of evaluations	% of evaluations	No. of evaluations	% of evaluations
Max points = 40 (Min required = 25)						
31 – 35	19	30.1%	36	49.3%	36	54.6%
25 – 30	10	15.9%	13 (1)	17.6%	15	22.7%
21 – 24	(1)	1.6%	(4)	5.5%		
16 – 20	-	-	(1)	1.4%		
11 – 15	-	-	(1)	1.4%		
6 – 10	-	-			-	-
Total	63 (1)	100.0%	73 (7)	100.0%	66	100.0%

Contributing Factors: Using Evaluation and Monitoring

Sub-Criterion 6.2 Effective Results Monitoring and Reporting

Almost all the evaluation reports with valid findings on sub-criterion 6.2 gave examples of weaknesses in monitoring and results reporting. Surprisingly, a full 15 evaluation reports found that monitoring systems were either absent or were not providing data on results. Another 15 evaluations reported that monitoring systems were in place but were weak, unreliable and lacked essential data. While there is no clear, over-riding explanation for the continuing problem of weak results monitoring and reporting there are some frequently noted limiting factors:

- Inadequate and inconsistent performance by partners charged with gathering and entering monitoring data, most often related to capacity issues among partner personnel (Zimbabwe, Malawi)
- Absence of harmonized data collection tools and/or agreed indicators among partners (Jordan, Guyana, Mali) and,
- Weak output and outcome indicators which were not well-defined during program development (Malawi).

In fact, the problem of poorly defined results definitions and resulting inadequate results frameworks and weak indicators is cited as the single most important contributor to ineffective monitoring systems and processes at UNICEF (18 cases). Without well-defined and explicit results and accompanying indicators it is impossible to establish meaningful results targets with verifiable indicators. This then leads to cascading problems of incomplete or missing data and a failure to compare measured outputs and outcomes to goals and targets.

Sub-Criterion 6.3 Effective RBM Systems

It is interesting that, despite the overall low level of positive results for the effectiveness of RBM systems (only 33 percent satisfactory or highly satisfactory), almost ten percent of the evaluations

reviewed in 2020 found these systems to be highly satisfactory. This is especially surprising given the relatively poor results for the strength of underlying monitoring systems (sub-criterion 6.2). Where RBM systems were found to be effective a few contributing factors which seem especially important:

- Program designs which included a detailed results framework. More importantly, the results frameworks included clearly defined outcomes with supporting indicators sometimes based on a clear program theory of change (Sudan, Congo, Benin, Korea)
- Regular and systematic use of the results reports by policy makers and managers (Kenya, Nepal, Croatia) and,
- Integration of RBM reporting into systems for performance-based financing (Zimbabwe).

Of course, RBM systems are reliant on reliable and regular inputs of results monitoring data so that all the negative factors noted for sub-criterion 6.3 on monitoring systems apply equally to the effectiveness of RBM systems. However, there are some negative contributing factors which are uniquely important in undermining the effectiveness of RBM systems at UNICEF.>

Sub-Criteria 6.1 Effective Evaluation Systems and 6.4 Using Evaluation to Improve Effectiveness

Where evaluation reports found evaluation systems and procedures to be effective, they most often pointed to the fact that evaluations were commissioned, planned, and carried out systematically and linked to the program life cycle (Congo, South Sudan, Kenya, Bangladesh, Egypt, Croatia). Others pointed to the independence of evaluations as a key strength (Zimbabwe, Guyana) and to effective use of baselines and reliance on participatory evaluations.

Finally, the improvement in the use of evaluation to improve effectiveness is rooted in the continuing improvement and completeness of the management response documents filed with the Evaluation Office and made public on the UNICEF evaluation web page. For the 2020 review all 62 evaluations in the sample were accompanied by a credible and thorough management response document.

5. Innovation

Summary – Innovation

Approximately half of the evaluation reports reviewed noted a significant effort by UNICEF to support some form of innovation in the program or activities being evaluated. In almost all cases this takes the form of either developing and implementing a new approach to program delivery or adapting an innovative good practice in programming from another jurisdiction. Overall, the level of innovation noted in the 2020 review matches the same finding in the 2016 review, which itself represented a major advance from 2012. These findings are consistent with those highlighted in the 2018 thematic evaluation of innovation at UNICEF.

“UNICEF has **clearly signaled its strategic intent** to utilize innovation as an important means to achieve results for children during the period 2014-2021. However, this strategic intent was not made sufficiently operational to set an organization- wide course or direction.

UNICEF lacks an explicit overarching ‘innovation strategy’, setting out how it intends to go about innovation, what it intends to achieve, and how staff throughout the organization can contribute. “¹⁹

The 2020 Development Effectiveness Review confirms that UNICEF continues to emphasize and invest resources in innovation. The main forms of innovation identified in the evaluations reviewed are the introduction of new approaches to program delivery or the introduction of an existing but innovative program delivery mechanism into a new jurisdiction.

5.1 Tracking Progress on Innovation

The criteria and sub-criteria used for all three effectiveness reviews at UNICEF (2012, 2016 and 2020) are derived directly from the evaluation criteria adopted and promoted by the Development Assistance Committee of the OECD. As a result, results presented for all the criteria (as in Section 4.0) do not deal with the role and extent of innovation at UNICEF.

Nonetheless, to reflect the importance of innovation as an important programming approach at UNICEF, the 2016 review began collecting and analyzing all references to innovation made in published UNICEF evaluation reports. This approach has been continued for the 2020 Review.

5.2 Trends and Progress in Innovation at UNICEF

The 2016 review noted 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 2016 review identified specific comments on technological innovations, new program models, or new implementation arrangements in 51 percent of reviewed evaluation reports²⁰.

The 2016 review also noted that UNICEF had succeeded in fostering a commitment to innovation in the programs evaluated from 2012 to 2015, although there was still scope for developing a stronger culture of innovation.

The current DER has confirmed that UNICEF evaluations published from 2016 to 2019 also often refer to significant innovations in programming approaches and methods and in new technologies for development. For the 2020 review 50 percent of evaluation reports refer to some form of innovation supported by UNICEF. This suggests that the frequent use of innovation in UNICEF programming has at least been maintained since 2016.

¹⁹ Evaluation of Innovation, p. 65.

²⁰ UNICEF, Development Effectiveness Review of United Nations Children Fund (UNICEF) - 2012-2015). P.61

In 2018, UNICEF published the: *Evaluation of Innovation in UNICEF Work - Synthesis Report* which included a three-part typology of innovation at UNICEF as a key element in its analytical approach:

- **Hard’ technologies**, commonly referred to as “product innovation” within UNICEF, to enable sector-specific results, with examples including diagnostic tools for the prevention of diseases, emergency shelters or water pumps
- **Information and communications technology (ICT)**, particularly mobile technologies, as cross-sectoral enablers for information management, service delivery, performance monitoring, participation, and advocacy
- **Innovative methodologies, approaches and processes** (‘soft’ technologies) like behaviour change approaches or design thinking, allowing for wider programmatic impact.²¹

The 2020 review indicates that the last of these three types – innovative programming methods, approaches and processes is by far the most frequently used form of innovation at UNICEF. Fully

Box 17: Types of Innovation Observed in 2020

1. Innovative Program Approaches (17 references)
2. Replication of an Existing Innovation in Programming (12 references)
3. Innovative Partnerships (2 references)
4. New Technologies Including ICT (2 references)

31 of the 32 evaluations noting the use of innovation referred to some form of innovative programming tool or “soft” technology.

As noted in box 15, the 2020 review identified four different types of innovations in UNICEF programming

referenced in reviewed evaluation reports (this is consistent with the approach taken in 2016). The most frequent use of innovation was either in the use of a new programming approach or in adapting an existing approach that was used in another context but represented an innovative application in a new region, country, or institution. A third form of innovation focused on innovative partnerships, including, for example, partnering with a private sector company to develop and Early Childhood Development (ECD) data base in Macedonia. There was only one example in the 62 evaluations reviewed of an effort by UNICEF to support the use of an innovative “hard technology” (Vietnam).

Examples of **innovations in programming** include:

- Use of “toy making” as an activity to increase availability of materials and engage parents in play and ECD (Uganda);
- Introducing a “step model” with graduated levels of UNICEF support to implementing partners with each step requiring more advanced skill sets and increased resources (South Sudan);
- Innovative use of geographic and vulnerable-group membership criteria for targeting out of school girls (Nigeria):

²¹ Evaluation of Innovation in UNICEF Work - Synthesis Report, 2018. P.10

- Innovative use of Learner Networking Cards to limit interruption to learning during seasonal migration by incorporating migration related decision-making mechanisms in pastoral communities (Ethiopia);
- Use of innovative models of positive parenting support (Croatia) and;
- Innovations in pedagogy for delivery of LSE (India).

Efforts to replicate innovative programming in new contexts and jurisdictions include:

- Use of clean health center competitions rolled out to national level (Mali);
- Working with faith-based leaders and peer to peer networks (Lebanon);
- Use of Rapid SMS to collect information on the reach of campaign interventions at ward level (Nigeria); and,
- Use of clusters as a crucial node in the cascading chain of training at school level (Zimbabwe).

Evaluation reports noted innovative partnership arrangements through the use of pooled funding managed by UNICEF in support of the health sector (Zimbabwe) and collaboration with a private company in Albania to develop and ECD data base at kindergarten level.

6. Evaluation of the Strategic Plan and Implications

Summary: Evaluation of the Strategic Plan 2018-2021

The findings of the 2020 effectiveness review do not provide direct evidence of the UNICEF contribution to progress toward achieving the 2030 agenda because they are based on findings from evaluation reports published during two Strategic Plans: reports which focus on programming largely developed and implemented before 2018. However, the review does provide evidence on how well the change strategies and organizational enablers emphasized in the current Strategic Plan match the organization strengths of UNICEF and the challenges it faces in 2020. The change strategies and organizational enablers of the current Strategic Plan can be seen as an effort to reinforce comparative advantages and to continue or strengthen investments in those areas where UNICEF has demonstrated a positive record of achieving results. The change strategies and enablers can also be seen as efforts to improve or strengthen areas of UNICEF performance where it had not yet demonstrated a strong track record such as action to address or mitigate climate change or leveraging the power of business. The 2020 Development Effectiveness Review provides evidence that UNICEF is well positioned to use most of the change strategies and organizational enablers to improve overall effectiveness during the 2018-2021 Strategic Plan.

6.1 Analytical Approach

As noted in Section 1.1, one purpose of the 2020 Development Effectiveness Review is to contribute to the evaluation of the UNICEF Strategic Plan 2018 – 2020. However, it is not possible to simply apply the results of the DER directly to the overall purpose of the SP Evaluation: “Identify lessons learned from the contribution of the Strategic Plan 2018-2021 to the achievement of the 2030 agenda, prior to the onset of COVID-19.”²² The fact that the effectiveness review covers the period 2016 to 2019 means that over half the evaluations in the sample were published under the preceding Strategic Plan (2014 to 2017) and focus on programming from that period. However, this does not mean that the findings of the review are, by definition, inapplicable to the SP Evaluation.

Rather, the findings reported in Section 4 above represent an assessment of the suitability of UNICEF as agency with an evolving history of program effectiveness measured by the DER to the demands and requirements of the Strategic Plan.

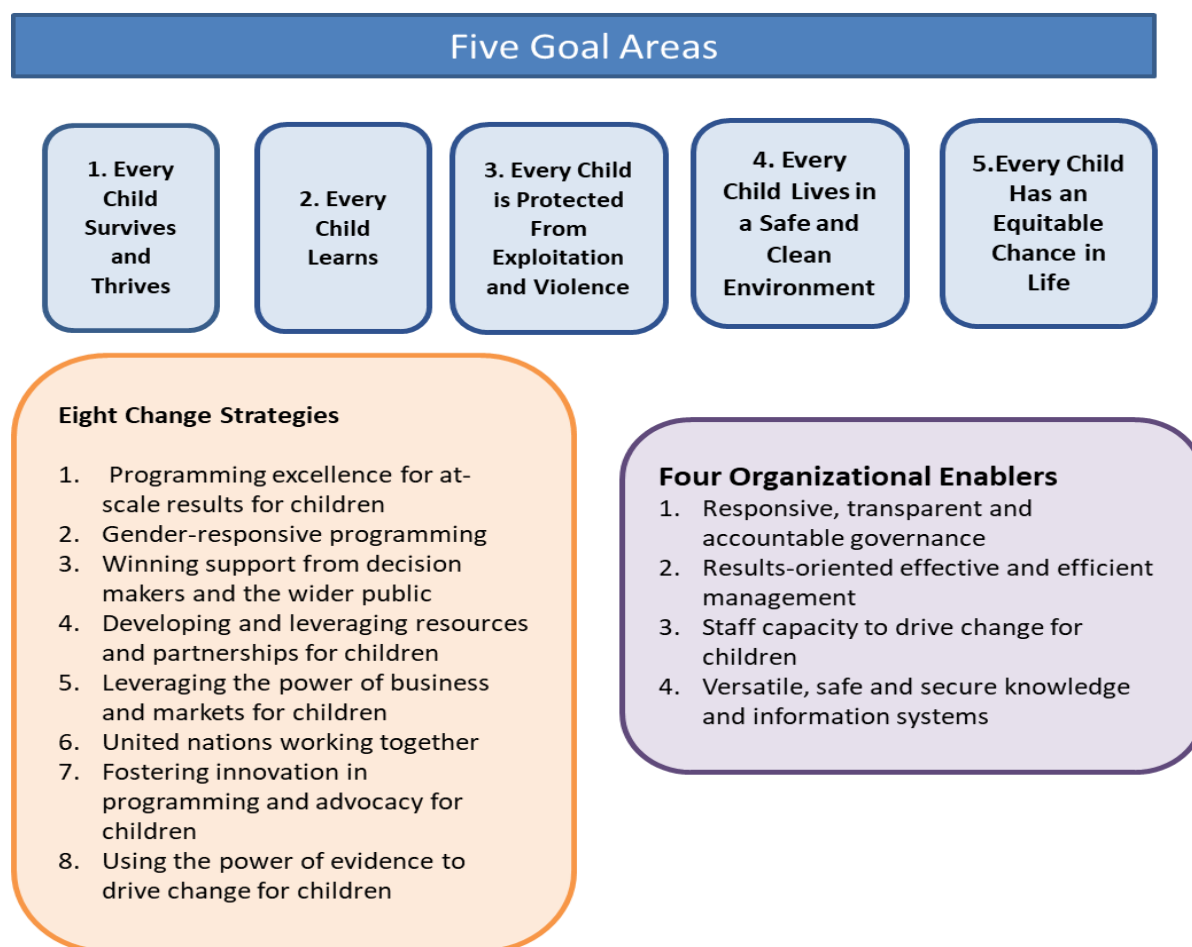
This approach fits well with the overall structure of the Strategic Plan 2018 – 2021 with its **five goal areas and accompanying results** which are to be achieved through the application of **eight change strategies** and four **organizational enablers** as illustrated in Figure 7.

The best evidence of the extent that UNICEF has been able to utilize the Strategic Plan 2018 – 2021 to contribute to the achievement of the SDGs and the 2030 agenda can be drawn first from the data gathered through the theory of change, results framework and monitoring system developed for the Strategic Plan and reported in the Mid-Term Review of the SP.

It seems clear this Development Effectiveness Review can best serve as an input to the SP Evaluation by providing an objective measure of how well UNICEF as an organization is suited to the identified change strategies and organizational enablers of the Strategic Plan. In other words, do the findings of the DER indicate that UNICEF can capitalize on the designated change strategies and organizational enablers? If the latter represent the foundations on which UNICEF intends to achieve results in the five goal areas, it is important to ask whether these are realistic elements of the Strategic Plan.

²² Draft Concept Note for the Evaluation of the UNICEF Strategic Plan 2018 - 2021

Figure 7: Goal Areas, Change Strategies, Organizational Enablers of the UNICEF Strategic Plan



The DER team built on this approach by cross matching the assessment sub-criteria used in the review to the Strategic Plan change strategies and enablers where they were most applicable. This relationship is illustrated in Table 4.

Table 12: Strategic Plan Change Strategies and Applicable Effectiveness Criteria

Change Strategies (CS) and Organizational Enablers (EN)	Applicable Effectiveness Assessment Sub-Criteria
CS 1: Programming at scale for children – programming that responds holistically to children’s needs and environments	1.1 Programs suited to needs of target group 2.1 Achievement of stated objectives – going to scale 2.3: Difference for substantial number of beneficiaries 4.2: Sustainability at institutional/community level 4.3: Strengthening enabling environment for development
CS 2: Gender-Responsive Programming – preventing child-marriage, GBV, and quality maternal, ante and post-natal care	3.1 Effective support to Gender Equity 3.3 Effectively addressing issues of equity

CS 3: Winning support – stronger advocacy, fundraising and communications force for children	1.3 Effective partnerships 2.4 Programs contribute to changes in national policies and programs 4.1: Sustainability
CS 4: Developing and leveraging resources and partnerships for children	1.3 Effective partnerships 4.1: Sustainability – benefits likely to continue – esp. factors including financial sustainability
CS 5: Leveraging the power of business and markets for children	1.3 Effective partnerships
CS 6: United Nations working together on common priorities	1.1 Programs meet the needs of target group members 1.3 Effective partnerships 3.3 Addressing equity
CS 7: Fostering innovation and advocacy for children	2.1 Achievement of stated objectives – contributing factors and role of innovation
CS 8: Using the power of evidence to drive change for children	6.4 Evaluation results used to improve development/humanitarian effectiveness
EN 1: More frequent data on results and resources and EN 2: Results oriented , effective, and efficient management	6.2 Effective systems for monitoring and reporting 6.3 RBM systems are effective

6.2 Evaluation Findings in Relation to the Strategic Plan

Building on the approach described in 6.1 above, the review team summarized the key findings of the 2020 Performance Review as they can be used to assess the suitability of change strategies and organizational enablers of the strategic plan. The results are summarized in table 5.

Table 13: DER Findings and Baseline for Change Strategies and Enablers

Change Strategies (CS) and Organizational Enablers (EN)	DER Findings Relevant to the Base-line for Change Strategies and Organizational Enablers at the Beginning of the 2018-2021 SP	Implications for the Evaluation of the Strategic Plan
<p>CS 1: Programming at scale for children – programming that responds holistically to children’s needs and environments</p>	<ul style="list-style-type: none"> 2.1 and 2.3. Objectives Achievement: UNICEF programs and activities achieve most of their stated objectives (75 percent of cases) and do result in benefits for large numbers of beneficiaries (95 percent) – often taken to scale. The extent that UNICEF supported programs are taken to scale and result in significant positive change for children and adolescents at a national level is one of the key findings of the review. When programs are not taken to scale, it is sometimes because UNICEF has supported pilot initiatives which are not accompanied by adequate unit cost analysis or are not integrated into national programs and budgets. An occasional negative factor is the absence of an adequately multi-sectoral approach to program design. 3.2 Cross Cutting Themes: Environmental Sustainability: UNICEF evaluations have not adequately addressed environmental sustainability. Where evaluations do take note of climate change it is mainly as an external contextual factor rather than an area for seeking results. 4.2. Sustainability: UNICEF has contributed to sustainable capacity development in both institutions and communities (73 percent). This area has seen steady improvement over the past ten years of programming (2012, 2016 and 2020 r partnerships. 2.4 Relevance: UNICEF has been effective in advocating for changes to national development plans and programs – including for disaster preparedness. This also relates to success in engaging in partnerships and the strength of UNICEF convening and coordinating powers. 	<ul style="list-style-type: none"> The DER finds that UNICEF supported programs were operating at scale and providing benefits to large numbers of children, however, under CS1 the SP should result in scaling at higher levels of targeted action to help achieve child-focused SDGs – especially given reported positive results on capacity building. Inadequate integration into national budgets is sometimes a problem for going to scale and CS1 could be expected to call for greater attention to this factor in Country Programs. The explicit reference to multi-sectoral approaches in CS1 represents an effort to address this issue in the SP – an area of need identified in evaluation reports. Evaluation studies commissioned before 2018 do not address climate change but the emphasis on climate mitigation in the SP could be expected to change this from 2019 onward to provide better evidence of contribution to results. Highly rated achievements in supporting institutional and community capacity development could lay the basis for more sustainable scaling of interventions. Represents an organizational comparative advantage for going to scale with policy changes and a concrete basis for acting on CS 3: Winning Support

Change Strategies (CS) and Organizational Enablers (EN)	DER Findings Relevant to the Base-line for Change Strategies and Organizational Enablers at the Beginning of the 2018-2021 SP	Implications for the Evaluation of the Strategic Plan
<p>CS 2: Gender-Responsive Programming – preventing child-marriage, GBV, and quality maternal, ante and post-natal care and empowering disadvantaged girls</p>	<ul style="list-style-type: none"> 3.1 Gender Equity: From a very low base in 2012, UNICEF has significantly improved the extent it engages in promoting Gender Equity and its success in contributing to results for GE (58 percent in 20 20). This reflects greater attention in the SP and, perhaps most importantly the adaptation of a specific Gender Action Plan (GAP) in 2016. 3.3 Contribution to Addressing Inequities: Most evaluations (67 percent) found UNICEF support effective in contributing to reduced inequities for children. Two sub-groups where UNICEF programming was seen as effectively addressing inequities were in programming for children with disabilities (noted in 30 percent of evaluations) and programming for adolescents (noted in 42 percent of evaluations in 2020). 	<ul style="list-style-type: none"> GE provides an example of how UNICEF can course-correct if sufficient managerial attention and resources are applied. Building on two successive Gender Action Plans, the results of the 2010 and 2016 Development Effectiveness Reviews, and (especially) the results of the recent GAP evaluation UNICEF has demonstrated effectiveness in addressing a cross-cutting priority such as gender (although the GAP Evaluation Report makes clear that more needs to be done). Up to the beginning of the SP, UNICEF evaluations most often identified positive results in addressing equity for children with disabilities and adolescents – the SP emphasis on disadvantaged girls should see broadening of results on addressing inequity beyond these two groups.
<p>CS 3: Winning support – stronger advocacy, fundraising and communications force for children</p>	<ul style="list-style-type: none"> 2.4 Relevance: UNICEF has been effective in advocating for changes to national development plans and programs – including for disaster preparedness. 4.1 Sustainability: sustainable financial support from national governments is an urgent challenge: Only 32 percent of evaluations found that benefits were likely to continue following program completion – most often because sustainable financing was not adequately addressed. 	<ul style="list-style-type: none"> UNICEF has demonstrated a track record of effectiveness in advocating for change for children, providing a strong basis for further strengthening this result in the SP An area where the SP could support strengthening effectiveness in advocacy concerns greater national budgetary commitment to sustain the benefits of UNICEF supported programs
<p>CS 4: Developing and leveraging resources and partnerships for children</p>	<ul style="list-style-type: none"> Sustainability: UNICEF has contributed to sustainable capacity development in both institutions and communities (73 percent). This area has seen steady improvement over the past ten years of programming (2012, 2016 and 2020 reviews). It relates in part to effective engagement in partnerships. 	<ul style="list-style-type: none"> The long-term trend to improved results in sustainable capacity development over more than ten years should, in turn, allow UNICEF to increase the level of leveraging of national resources to benefit children.

Change Strategies (CS) and Organizational Enablers (EN)	DER Findings Relevant to the Base-line for Change Strategies and Organizational Enablers at the Beginning of the 2018-2021 SP	Implications for the Evaluation of the Strategic Plan
	<ul style="list-style-type: none"> 2.4: Sustainability UNICEF has been effective in advocating for changes to national development plans and programs – including for disaster preparedness. Also relates to success in engaging in partnerships – convening and coordinating powers. 	<ul style="list-style-type: none"> DER findings on effective advocacy and the use of convening and coordinating power by UNICEF indicate there is significant scope to leverage resources as an SP priority.
CS 5: Leveraging the power of business and markets for children	<ul style="list-style-type: none"> Evaluations commissioned before the current SP provide few examples of successful efforts to engage with business/private sector at international/national/local level 	<ul style="list-style-type: none"> CS5 was not a strategic priority for UNICEF prior to 2018 and, as a result, does not figure in evaluations reviewed. The advent of programmatic action such as Business 4 Results (B4R) at country level and continuing senior management support during the SP are direct efforts to address this issue.
CS 6: United Nations working together on common priorities:	<ul style="list-style-type: none"> Relevance: UNICEF has demonstrated a strong capacity to engage in and support partnerships within the UN and with national government agencies, national NGOs, international NGOs, development partners and CSOs (85 percent). This helps programs to adjust to changing priorities. Evaluations have identified UNICEF contributions to five of the six joint UN priorities identified for this change strategy. Action to address climate change is the one priority area which is not identified in any of the sampled evaluations. 	<ul style="list-style-type: none"> Evaluations indicate that UNICEF has been largely supportive of efforts to coordinate and collaborate with UN partners at country level in keeping with the work of UNCTs and instruments such as the UN Partnership Framework – providing a basis for continuing work on joint priorities As already noted, climate change mitigation is a new cross-cutting result for UNICEF and the current SP provides a platform for raising visibility on climate change. Evaluations commissioned during the SP should provide more evidence in this area.
CS 7: Fostering innovation and advocacy for children	<ul style="list-style-type: none"> UNICEF evaluations are increasingly likely to identify innovation as a factor in achieving positive results. Over half the evaluations reviewed in 2020 identified innovations (54 percent) identified positive innovations – most frequently related to innovative programming approaches but also including technology innovations. A major increase since 2016. The 2018 Evaluation of Innovation work at UNICEF noted that UNICEF had clearly signaled its strategic intent to utilize innovation but that it lacked an overarching innovation strategy. 	<ul style="list-style-type: none"> Identified gaps in fostering innovation which could be addressed by the SP include; <ul style="list-style-type: none"> - The absence of an explicit, overarching innovation strategy - Limited use of innovation in a number of Country Offices reflecting a need for better communication/encouragement to CO staff. - Poor understanding among funders of the UNICEF strategy for innovation which hinders

Change Strategies (CS) and Organizational Enablers (EN)	DER Findings Relevant to the Base-line for Change Strategies and Organizational Enablers at the Beginning of the 2018-2021 SP	Implications for the Evaluation of the Strategic Plan
		efforts to mobilize resources for innovative programming
<p>CS 8: Using the power of evidence to drive change for children</p> <p>EN 1: More frequent data on results and resources and</p> <p>EN 2: Results oriented, effective, and efficient management</p>	<ul style="list-style-type: none"> 6.2 Monitoring and Evaluation: Results monitoring and reporting systems remain weak with 41 percent of evaluations with positive results in 2020 compared with 39 percent in 2012. Lack of baseline data and implementing partner capacities are important constraints. 6.3 Monitoring and Evaluation: At least as reported in evaluations published by UNICEF to the end of 2019, the promise of RBM largely unfulfilled. In the 2020 review only 33 percent of cases reported effective systems for RBM with no noticeable improvement over 2016. However, evaluations systems are found effective (6.1) in 77 percent of cases and evaluation results are used to improve program effectiveness in 96 percent of 2020 evaluations. 	<ul style="list-style-type: none"> The 2018-2021 SP, was the first to roll-out a comprehensive Results Framework with detailed results definitions, indicators, baselines and targets at the goal, outcome and output levels. This should contribute to improved findings in evaluations commissioned from 2019 onward. The apparent task for the SP (post 2019 MTR) is to better link results across the output and outcome levels while managing the reporting burden for Country Offices. Strengthened results reporting and the continuing trend to improved evaluation systems and processes at program level could allow UNICEF to strengthen the “contribution story” required to support resource mobilization.

Table 13 highlights the areas where the findings reported in Section 4.0 may be most useful in evaluating specific elements of the Strategic Plan 2018 – 2021. During the upcoming evaluation of the Strategic Plan, these will be further investigated and elaborated.

7. Conclusions

7.1 Conclusions

Relevance

1. UNICEF has **consistently demonstrated a high level of relevance** of the programs and activities it supports over a ten-year period encompassing all three effectiveness reviews (2012, 2016 and 2020). UNICEF supported programs are **well suited** to the needs and priorities of the children and adolescents they target, remain consistent with national priorities, and often provide benefits at a national scale. There is a **clear link from careful needs analysis during program design** to programs that are **well suited to the needs and priorities of children and adolescents**. UNICEF can strengthen this link by continuing/increasing priority and **emphasis given to consultation with** children, adolescents, and other **community stakeholders** during program development.
2. UNICEF has demonstrated a **strong commitment to engage in and support partnerships with a wide range of stakeholders** including national government agencies, civil society organizations, UN Country Team (UNCT) members and other bilateral and multilateral development partners. This represents an important organizational comparative advantage.
3. Building on strong partnerships with governments and other stakeholders, UNICEF has demonstrated a **capacity to support programming taken to scale at a national level**. This combined with the fact that programs suit the needs of children and adolescents provides evidence of a strong positive contribution to outcomes at a national level – including progress toward child-relevant SDGs.

Objectives Achievement

4. UNICEF has maintained a **strong record of positive program performance in objectives achievement** across all three effectiveness reviews. This finding is reinforced by the positive results of the 2020 Development Effectiveness Review. Almost all UNICEF supported programs and activities (95 percent) provide significant benefits to target group members. However, **there is scope to improve targeting** to provide more direct benefit from UNICEF supported programs to marginalized children and adolescents.
5. UNICEF has made an **effective contribution to the development of national policies, plans and programs for children** as measured across all three effectiveness reviews (2012,2016 and 2020). In recent years (2016 to 2019), UNICEF evaluations have focused more closely on the operational effectiveness of programs than on “upstream” effects on policy. This represents a **shift in the focus of evaluations rather than a reduction in the level of policy advocacy by UNICEF**.

Cross-Cutting Themes: Gender Equity, Environment, Promoting Equity

6. UNICEF supported programs continue to improve in the level of attention and the positive results achieved to improve Gender Equity. This is reflective of **increased attention to Gender Equity under successive Gender Action Plans**. In addition, UNICEF **evaluations are addressing Gender Equity more consistently** and with more in-depth analysis. Further, the improvement in results for contributing to Gender Equity demonstrate an organizational ability to successfully implement strategies for change at UNICEF. Nonetheless there remains an opportunity to build on the steady progress on Gender Equity made during the 2016 to 2019 period and to raise **the overall level of organizational ambition with regard to gender equity** as recognized by the recent evaluation of the GAP.
7. UNICEF evaluations **do not generally address contributions to environmental sustainability in any systematic way**. The reviewed evaluation reports made **very little reference to efforts to address or mitigate the effects of climate change** in the programs and activities UNICEF supports. This **lack of attention** to the broader positive or negative environmental effects of UNICEF supported activities, including those related to climate changes, **weakens UNICEF's ability to develop and support a credible contribution story regarding climate change**: despite a strong organizational track record in, for example, addressing water supply insecurity in climate change affected regions and countries. **Evaluation and monitoring systems** and practices at UNICEF will **need to clearly address issues of environmental impact and climate change mitigation** to strengthen legitimate claims of UNICEF contribution.
8. The 2020 effectiveness review reports **positive results for UNICEF efforts to advance equity for marginalized and vulnerable children and adolescents** – most often with program elements targeting **adolescents and children with disabilities** – a reflection of recent organizational emphasis on these two groups of disadvantaged. This is another area where **increased emphasis on meaningful engagement and consultation with community members**, especially vulnerable children and adolescents and their families during program development **can improve UNICEF effectiveness**. This would assist UNICEF in extending the application of the principles leave no-one behind to a wider set of the most vulnerable children and adolescents.

Sustainability

9. The **sustainability of the benefits of UNICEF supported programs, especially financial sustainability remains a critical challenge**. Experience with programs that have proven sustainable indicates that responding more effectively to this challenge will require: intensified pre-program **planning for resource mobilization** from non-UNICEF sources on program completion; ensuring that **pilot, locally-scaled programs are designed with adequate plans for going to scale**, including cost effectiveness data gathering; and ensuring that UNICEF supported programs include realistic **exit strategies** for the period when UNICEF support is being withdrawn. While UNICEF has been generally successful in supporting programs delivered at scale, evaluation reports indicate

their medium to longer term financial viability is a continuing and significant challenge. There is also a need to **ensure that UNICEF supported programs are embedded in national budgets** with access to sustainable financing from **governments, the private sector and civil society**. SP Change Strategy 5 – Leveraging the Power of Business and Markets for Children may provide an important avenue to improving the sustainability of at-scale interventions and programs supported by UNICEF.

10. UNICEF has been **successful in developing the capacity of partner institutions and communities** so that they can continue with the delivery of services following program completion (provided financing can be secured). This reflects both a consistent investment in planning for capacity development and a long history of effective partnership at UNICEF.

Efficiency

11. **Efficiency and value-for-money issues are receiving more attention in recently published UNICEF evaluations.** In addition, the 2020 review indicates that **more UNICEF offices and staff are demonstrating a “cost-effectiveness mind set”** in developing and supporting programs. However, **cost and resource-efficiency remain a major challenge.** UNICEF supported programs can **also encounter considerable delay in implementation** due to over-ambitious timelines, administrative delays in approvals, problems in disbursement and issues in supply-chain management. These findings have been consistent across all three effectiveness reviews and have shown little to no improvement over ten years of programming. The **current initiative to re-examine UNICEF business models** and operational practices in different country contexts (such as income category) will need to address questions relating to **improved business practices for more timely and efficient program delivery.**

Using Monitoring and Evaluation to Improve Effectiveness

12. UNICEF **systems for evaluation have steadily improved in overall effectiveness over ten years of programming** and across all three effectiveness reviews. This improvement is supported by the concurrent improvement in the quality scores of evaluation reports assessed in the 2012, 2016, and 2020 effectiveness reviews. In addition, the number, responsiveness and **quality of the management response to evaluation reports have significantly improved.**
13. The **UNICEF Strategic Plan, 2018 -2021**, included a **comprehensive Results Framework with results, related performance targets, and indicators** developed at the Goal, Outcome and Output level. Wherever possible, especially at output level, the results framework also includes baseline data. **Over time this should improve evaluation findings related to the strength of monitoring and RBM systems at UNICEF.** However, evaluations reports reviewed in 2020 indicate that **the effectiveness of monitoring and RBM systems at country program level remains a critically important challenge.** Often the absence of a credible baseline and valid change indicators reported against time-bound targets frustrated efforts at results reporting and RBM. If UNICEF is to improve the visibility and credibility of claims regarding its contribution to

changes at country level it will need to **better link the baselines, targets and results of the SP Results Framework to program monitoring systems in operation at country level.**

Innovation

14. UNICEF supported programs and activities reflect a **significant effort to support innovation**. Innovations **in program approach**, including the application of new models to service delivery **are more frequent** in UNICEF supported programs than “hard innovations” such as new information and communications technology. UNICEF has signaled its strategic intent to use innovation in the current SP but there is still considerable room **for more strategic and systematic use of innovation** in at least half of the programs covered by evaluations in the 2020 review. As demonstrated by the relative success of the Gender Action Plans and as pointed out in the recently published thematic evaluation of innovation, **greater effectiveness through innovation** may require the development of “an explicit overarching **innovation strategy**”, a **better understanding of staff capacity** for innovation, **better communications** on innovation as a priority and **improved accountability** for innovation.

Evaluating the Strategic Plan 2018-2020

15. The Development Effectiveness Review 2020 provides **useful evidence** regarding **base-line conditions** prevailing at the beginning of the Strategic Plan for some of the most important **change strategies (CS) and organizational enablers (EN)**. In particular, the findings of the review support the thesis that **CS1: Programming at Scale** for Children represents an important comparative advantage for UNICEF based on the relevance and scale of UNICEF programming and its track record on effective partnerships. However, inadequate integration into national budgets for some programs suggest that **greater attention to securing national funding** could improve effectiveness in both going to scale and sustaining program benefits.
16. The results of the 2020 review indicate that **UNICEF is well positioned to capitalize on CS2: Gender-Responsive Programming**. The mainly successful implementation of two sequential Gender Action Plans (2014-17 and 2018-21) provides important evidence that UNICEF can **be successful by making significant investments in organizational architecture** (including staff capacities and programming guidelines) to achieve **changes at a strategic and operational level in pursuit of clear priorities**. It also indicates that there is continuing scope for increased UNICEF ambition and attention to Gender Equity which **further raises the importance of CS2**. This is an important lesson for the Strategic Plan. It suggests that continued investment of time, attention and resources to CS5 **Leveraging the Power of Business** can be an effective change strategy.
17. Regarding **CS 3: Winning support through stronger advocacy**, fundraising and communications for children, UNICEF has demonstrated a **strong track record of effectiveness in advocating for changed policies and programs** to the benefit of children, especially at national level. Evaluation reports, rarely address the question of

fundraising, however, they do indicate that **securing national budgetary commitments is critical to sustaining the benefits of UNICEF supporting programs.**

18. UNICEF has contributed to a **long-term trend in improved results in sustainable capacity development with, in particular, national institutions of governance.** This should allow **UNICEF to build on CS4:** Developing and leveraging resources and partnerships for children by **increasing the leveraging of national resources.** UNICEF can also rely on demonstrated effectiveness in advocacy and the use of its convening and coordinating powers to provide scope for leveraging resources.
19. **CS5, Leveraging the power of business and markets, is a relatively new strategic priority** for UNICEF, introduced in the Strategic Plan 2018-2021. Not surprisingly, it has not been examined in the evaluation reports reviewed in 2020 since almost all were commissioned before 2019. The advent of programmatic action such as the Business 4 Results (B4R) initiative and continued senior management support represent actions taken since 2018 to act on CS5. There is **an opportunity for UNICEF to study recent experience** in engaging with businesses on common objectives to identify and reduce barriers to effective cooperation and to **identify good practices and refine guidelines** so that country offices can confidently implement CS 5.
20. UNICEF has been **effective in applying Change Strategy 6: United Nations working together** based mainly on a long-standing commitment to **partnership with UNCT members and national governments.** It has also demonstrated the capacity to implement Change Strategy 7 – Fostering Innovation. However, up to 2018 at least, UNICEF lacked a fully developed plan and strategy for supporting innovation.
21. UNICEF has taken action to strengthen results reporting at the Strategic Plan level with the establishment of the DAPM and its efforts to improve country office reporting. In addition, the 2018-2021 SP, was the first to roll-out a comprehensive Results Framework with detailed results definitions, indicators, baselines and targets at the goal, outcome and output levels. This should, over time, help UNICEF to act more effectively in implementing **CS8: Using the power of evidence to drive change for children** and **EN 2: Results oriented, effective and efficient management.** This would help UNICEF to strengthen the credibility of the “contribution story” it makes regarding results at the goal, outcome and output levels.
22. Building on the SP Results Framework to effectively realize **EN2: Results oriented management** will require UNICEF to: a) **better link baselines, targets, and results data gathered for SP reporting requirements to monitoring systems at country program level** and b) **better link results reported at the output level** (where UNICEF has more control and accountability) **to results at outcome and goal level** where contribution and accountability are shared among many different actors.

Annex A: Criteria for Assessing Developmental and Humanitarian Effectiveness

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

Annex B: Evaluation Sample (62 evaluations)

Title	Year	Region	Country	Country Income Level	Thematic Area	Humanitarian crisis	E M R	GEROS Rating	Language
2019 Kyrgyzstan: Summative evaluation of implementation of the State Program on Justice for Children in Kyrgyzstan 2014-2018	2019	ECA	Kyrgyzstan	Lower-Middle	Child Protection	3	1	2	E
2019 Malawi: Evaluation of the Community Led Total Sanitation and Hygiene Program Phase I: (Qualitative and Secondary Quantitative Analysis)	2019	ESAR	Malawi	Low	WASH	1	1	N/A	E
2018 Romania: "School attendance initiative" - School participation campaign implemented by UNICEF Romania	2018	ECA	Romania	Upper-Middle	Education	3	1	2	E
2018 Kazakhstan: Evaluation of the Justice for Children Reform in Kazakhstan	2018	ECA	Kazakhstan	Upper-Middle	Child Protection	3	1	2	E
2018 Vietnam: External After Action Review of the Government of Viet Nam and UNICEF emergency response supporting children and women in 10 provinces in Viet Nam affected by drought and salt water intrusion crisis (2016-2017)	2018	EAPR	Vietnam	Lower-Middle	Nutrition, WASH, Humanitarian	3	1	2	E
2018 Thailand: An Evaluation of Advocacy as a Strategy in the UNICEF Thailand Country Program 2012-2016	2018	EAPR	Thailand	Upper-Middle	Child Protection, Social Inclusion	3	1	2	E
2018 Korea: Evaluation of Community-based Management of Acute Malnutrition Program supported by UNICEF in DPR Korea 2015–2017	2018	EAPR	DPR Korea	Low	Nutrition	3	1	1	E
<u>2018 Zimbabwe: Summative Evaluation of UNICEF support for Education in Zimbabwe</u>	2018	ESAR	Zimbabwe	Lower-Middle	Education	3	1	1	E
2018 Malawi: Evaluation of Malawi Child Protection Strategy 2012-2018	2018	ESAR	Malawi	Low	Child Protection	3	1	2	E
2018 Kenya: Evaluation of NICHE in the first 1,000 days of a child's life in Kitui and Machakos counties, Kenya	2018	ESAR	Kenya	Lower-Middle	Nutrition	1	1	3	E
2018 Paraguay: Evaluación de procesos del Programa de supervivencia infantil y materna de Alto Paraná y Canindeyú 2014-2019. LRPS-9134651- 2017	2018	LACR	Paraguay	Upper-Middle	Health	3	1	2	S
2018 Guyana and Suriname: Evaluation of the Guyana nursery education program	2018	LACR	Guyana, Suriname	Upper-Middle	Education	3	1	2	E
2018 State of Palestine: Evaluation of Family Centres in Gaza	2018	MENA	State of Palestine (West Bank and Gaza)	Lower-Middle	Health, Education, Child Protection, Social Inclusion, Humanitarian	3	1	1	E
2018 Lebanon: Evaluation of Immunisation Program (EPI) in Lebanon (2013-2017)	2018	MENA	Lebanon	Upper-Middle	Health	3	1	2	E

Title	Year	Region	Country	Country Income Level	Thematic Area	Humanitarian crisis	E M R	GEROS Rating	Language
2018 Pakistan: Evaluation of Mother and Child Weeks Program, Pakistan	2018	ROSA	Pakistan	Lower-Middle	Health	3	1	2	E
2018 India: Evaluation of Adolescent Life Skills Education Program in Maharashtra	2018	ROSA	India	Lower-Middle	Education	3	1	2	E
2018 Togo: Evaluation Sommativ de l'UNICEF - Togo avec les radios pour le changement de comportement et l'adoption de pratiques favorables au bien-etre des enfants (2012-2018)	2018	WCAR	Togo	Low	Health, Education, Child Protection, Social Inclusion	3	1	2	F
2018 Niger: Evaluation des cliniques mobiles dans la région de Diffa	2018	WCAR	Niger	Low	Health, Humanitarian	1	1	2	F
2018 Mali: Évaluation sommative de l'approche de Programmation intégrée dans les régions de Mopti et Sikasso au Mali	2018	WCAR	Mali	Low	Health, Nutrition, WASH	3	1	2	F
2018 Ghana: Summative Evaluation of Ghana's Child Protection System Strengthening at District Level	2018	WCAR	Ghana	Lower-Middle	Education, Child Protection	3	1	2	E
2018 Benin: Evaluation finale du projet « Libérer le potentiel des enfants » au Benin Période : 2014 à 2016	2018	WCAR	Benin	Low	Education, Social Inclusion	3	1	2	F
2017 BandH: Evaluation of the UNICEF-supported Component of the Project "Support for Durable Solutions of the Revised Strategy for Implementation of Annex VII of the Dayton Peace Agreement, in Bosnia and Herzegovina	2017	ECA	Bosnia and Herzegovina	Upper-Middle	Social Inclusion	3	1	1	E
2017 Albania: Evaluation of the "Breaking the cycle of exclusion for Roma children through Early Childhood Development and Education" multi-country project in the Former Yugoslav Republic of Macedonia, Serbia and Albania	2017	ECA	Albania	Upper-Middle	Education	3	1	2	E
2017 Indonesia: Monitoring and Evaluation of PKH Prestasi Pilot Project	2017	EAPR	Indonesia	Lower-Middle	Nutrition	3	1	3	E
2017 Somalia: Evaluation of Social Mobilization Network (SMNet)	2017	ESAR	Somalia	Low	Health	1	1	3	E
2017 Ethiopia: An Impact Evaluation of Alternative Basic Education in Ethiopia	2017	ESAR	Ethiopia	Low	Education	1	1	2	E
2017 Honduras: Evaluación de implementación estrategia Retorno de la Alegría para la niñez migrante en Honduras	2017	LACR	Honduras	Lower-Middle	Child Protection, Social Inclusion	3	1	2	S
2017 Lebanon: Evaluation of the Water, Sanitation and Hygiene (WASH) Program within the UNICEF Country Program in Lebanon (2013-2016)	2017	MENA	Lebanon	Upper-Middle	WASH	3	1	2	E

Title	Year	Region	Country	Country Income Level	Thematic Area	Humanitarian crisis	E M R	GEROS Rating	Language
2017 Egypt: The Evaluation of Meshwary Project Phase II	2017	MENA	Egypt	Lower-Middle	Education, Social Inclusion	3	1	2	E
2017 Bangladesh: Program Evaluation of UNICEF Bangladesh Communication for Development (C4D) Program from 2012 to 2016	2017	ROSA	Bangladesh	Lower-Middle	Health, Education, Child Protection, WASH, Social Inclusion, Gender Equity	2	1	1	E
2017 Republic of Cameroon: EVALUATION DU PROGRAM WASH UNICEF CAMEROUN 2013-2016	2017	WCAR	Republic of Cameroon	Lower-Middle	WASH	1	1	2	F
2017 Mali: Evaluation du projet sur la promotion d'EAH dans les structures de soin au Mali	2017	WCAR	Mali	Low	WASH	3	1	2	F
2017 Guinea Bissau: Final Evaluation of the Community Health Component of the "Program for reducing Maternal and Infant Mortality (PIMI) in Guinea-Bissau	2017	WCAR	Guinea Bissau	Low	Health	3	1	2	E
2017 Congo: Évaluation des interventions des nations unies en faveur des réfugiés au Congo	2017	WCAR	Congo	Lower-Middle	Health, Education, Child Protection, WASH, Humanitarian	3	1	2	F
2016 Croatia: Evaluation of Parenting Support Programs	2016	ECA	Croatia	High	Social Inclusion	3	1	2	E
2016 Myanmar: Evaluation of QBEPs Township Education Improvement Plan Activities	2016	EAPR	Myanmar	Lower-Middle	Education	3	1	1	E
2016 Fiji: Evaluation of Capacity Development for Police on Child Protection in Fiji, Kiribati and Vanuatu	2016	EAPR	Fiji	Upper-Middle	Child Protection	3	1	3	E
2016 South Sudan: Multi-Country Real Time Evaluation of UNICEF Gender-Based Violence in Emergencies Programs: South Sudan Country Report	2016	ESAR	South Sudan	Low	Gender Equity, Humanitarian	1	1	2	E
2016 Rwanda: Rwanda RapidSMS Impact Evaluation	2016	ESAR	Rwanda	Low	Health	3	1	2	E
2016 Malawi: Malawi Social Cash Transfer Program Endline Impact Evaluation Report	2016	ESAR	Malawi	Low	Social Inclusion	3	1	3	E
2016 Kenya: Evaluation of the Integrated Intervention within the CFS framework	2016	ESAR	Kenya	Lower-Middle	Education	3	1	3	E
2016 Ethiopia: Evaluation of Mobile health teams in Afar and Somali Regions	2016	ESAR	Ethiopia	Low	Health	3	1	2	E
2016 Haiti: Evaluation de Kore Fanmi	2016	LACR	Haiti	Low	Health, Nutrition, Education, Child Protection, Social	1	1	2	F

Title	Year	Region	Country	Country Income Level	Thematic Area	Humanitarian crisis	E M R	GEROS Rating	Language
					Inclusion, Gender Equity				
2016 Argentina: Evaluation of the Program ICTmediated secondary rural schools	2016	LACR	Argentina	Upper-Middle	Education	3	1	2	S
2016 Jordan: Evaluation of Infant and Young Child Feeding Program	2016	MENA	Jordan	Upper-Middle	Health, Nutrition	3	1	2	E
2016 Algeria: Evaluation du projet sur la santé des mères et des nouveau-nés dans les camps de réfugiés sahraouis de Tindouf	2016	MENA	Algeria	Upper-Middle	Health, Humanitarian	3	1	3	F
2016 Nepal: Evaluation of UNICEF's Response and Recovery Efforts to the Gorkha Earthquake in Nepal (25 April 2015-31 January 2016)	2016	ROSA	Nepal	Low	Humanitarian	3	1	2	E
2016 India: Evaluation of Pilot DHaAL Interventions for Group Handwashing with Soap Before Mid-Day Meals in Rural Elementary Schools of Assam 2013-14	2016	ROSA	India	Lower-Middle	WASH	3	1	2	E
2016 Zimbabwe: Independent Evaluation of the Health Transition Fund in Zimbabwe	2016	WCAR	Zimbabwe	Lower-Middle	Health, HIV/AIDS	3	1	2	E
2016 Nigeria: Evaluation of the Maternal, Newborn and Child Health Week in Nigeria	2016	WCAR	Nigeria	Lower-Middle	Health	3	1	2	E
2016 Guinea Bissau: Evaluation of the UNICEF Supported Adult Literacy Program in Guinea-Bissau	2016	WCAR	Guinea Bissau	Low	Education, Gender Equity	3	1	3	E
2016 Burkina Faso: Evaluation finale du projet de « renforcement des services d'hygiène et d'assainissement de base et amélioration des pratiques d'hygiène nutritionnelle en milieu familial dans 8 régions du Burkina Faso	2016	WCAR	Burkina Faso	Low	WASH	3	1	2	F
2019 Uganda: Early Childhood Development Kit Humanitarian Evaluation	2019	ESAR	Uganda	Low	Health, Humanitarian	1	1	N/A	E
2019 Burundi: Evaluation des campagnes "Back To School" 2016, 2017 et 2018 au Burundi	2019	ESAR	Burundi	Low	Education, Social Inclusion	1	1	N/A	F
2018 Jordan: Evaluation of the UNICEF-supported specialized child protection case management response	2018	MENA	Jordan	Upper-Middle	Child Protection, Social Inclusion, Humanitarian	3	1	N/A	E
2017 State of Palestine: Evaluation for Humanitarian Action for Children	2017	MENA	State of Palestine (West Bank and Gaza)	Lower-Middle	Health, Education, Child Protection, WASH, Humanitarian	3	1	2	E
2017 Lebanon: Evaluation of the UNICEF Child Protection Program for Vulnerable Children and Women in Lebanon 2013-2016	2017	MENA	Lebanon	Upper-Middle	Child Protection, Social Inclusion, Humanitarian	3	1	2	E

Title	Year	Region	Country	Country Income Level	Thematic Area	Humanitarian crisis	EMR	GEROS Rating	Language
2016 Sudan: Evaluation of Humanitarian Action: Child Survival in North Darfur	2016	MENA	Sudan	Lower-Middle	Health, WASH, Humanitarian	3	1	2	E
2016 Namibia: Evaluation of Namibia's PMTCT Program	2016	ESAR	Namibia	Upper-Middle	HIV/AIDS	3	1	2	E
2017 Nigeria: Impact Evaluation of UNICEF Nigeria Girls' Education Project Phase 3 (GEP3) Cash Transfer Program (CTP) in Niger and Sokoto States	2017	WCAR	Nigeria	Lower-Middle	Education, Social Inclusion, Gender Equality	2	1	2	E
2017 Eswatini: EVALUATION OF THE SWAZILAND CHILD FRIENDLY SCHOOLS (CFS) PROGRAM	2017	ESAR	Eswatini	Lower-Middle	Education	3	1	2	E
<u>2018 Benin: Evaluation du Programme de Cours Accélérés (PCA) au Bénin Période 2011 à 2018</u>	2018	WCAR	Benin	Low	Education	3	1	2	F

Legend:

- Region:** ECA – Eastern Europe and Central Asia; ESAR – Eastern and Southern Africa; EAPR – Eastern Asia and the Pacific; LACR – Latin American and the Caribbean; MENA – Middle-East and Northern Africa; ROSA – South Asia; WCAR – Western and Central Africa
- EMR (Evaluation Management Response):** 1 – Yes; 2 - No
- GEROS Rating:** 1 – Highly Satisfactory; 2 –Satisfactory; 3 – Fair; 4 Unsatisfactory
- Language:** - E – English; F – French; S – Spanish

Annex C: Methodology

The Evaluation Population and Sample

The process used by the review team to select a sample of evaluation reports in 2020 was the same as the process used in previous DERs. The following sampling steps were carried out by the review team:

1. Identified the overall universe of UNICEF evaluations published from 2016 to 2019 and profile those evaluations by year of publication, type of evaluation study, (program or intervention evaluation, case study report, real-time evaluation, mid-term evaluation), region or country focus (including global), thematic area of programming, and (new for 2020) UNICEF Strategic Plan Goal Area covered.
2. Eliminated those evaluation reports not suitable for the quantitative, meta-analysis component of a Development Effectiveness Review.
 - Global evaluations which may contain country program evaluations and would involve double counting of summarized results;
 - Regional evaluations for the same reasons as global evaluations;
 - Mid-term and Formative evaluations which report interim rather than final evaluation results and which could be superseded by subsequent end-line evaluations;
 - Studies, reviews and evaluability assessments which would not include evaluation findings;
 - UNDAF and UN Joint programming evaluations where results are not attributable to UNICEF programming;
 - Thematic evaluations at the country level;
 - Real-time evaluations (3) as prior DER exercises have shown that RTEs rarely address OECD/DAC developmental effectiveness criteria;
 - Case studies (9) which lack evaluation findings; and,
 - Evaluation reports which have been reviewed and received a GEROS rating of less than three (fair).

Note: Important global and thematic evaluations not covered in the quantitative review were qualitatively reviewed by the review team leader and deputy team leader with the results of that review being incorporated into the overall report to provide a global context for the quantitative findings.

3. Selected a preliminary sample.

The remaining body of published UNICEF evaluation reports (222 reports) then became the universe of evaluation reports for sampling purposes for the 2020 DER. This is the

step in the process where **the methodology introduced an element of randomization**. Evaluation reports were randomly selected from the universe of reports by beginning with a randomly selected number (in this case 1) and then selecting every third and fourth evaluation report. This resulted in a preliminary sample of 66 evaluation reports. Based on past experience, it is likely that a small number of reports may not be suitable for review due to lower quality or to double counting (for example, the same report may be published under different titles). By the end of the review process, there will be a minimum of 60 evaluation reports quantified and analyzed (not counting the qualitative review of thematic and global evaluations). As a result, the quantitative review of results of published evaluations **covered 62 evaluation reports or 28 percent** of the universe of suitable reports.²³

4. Compared the preliminary sample to the universe of published UNICEF evaluations and to the pattern of recent UNICEF expenditures by region, thematic area, and goal area to determine whether or not it is (reasonably) illustrative of UNICEF programming.
5. Adjusted the preliminary sample to improve the fit between the sample of evaluation reports and reviewed the final evaluation sample with UNICEF Evaluation Office.

When compared to the overall pattern of UNICEF Expenditures, (tables 11 to 13 in Annex 4) the preliminary sample **had some imbalances by region:**

1. **Over-representation of evaluations** focusing on Europe and Central Asia (15 percent of the sample versus 4 percent of expenditures in 2018) and evaluations focusing on Latin America and the Caribbean (15 percent of the sample versus three percent of 2018 expenditures); and,
2. **Under-representation of evaluations focusing on programs in the Middle East and North Africa region** (9 percent of the sample versus 25 percent of expenditures in 2018).

This imbalance arose mainly because the volume of evaluation reports published by UNICEF from 2016 to 2019 under-represents the volume of expenditures in MENA in the same period (just nine percent of evaluation reports when compared to 25 percent of expenditures). To redress this regional imbalance in the sample, the review team **developed the final evaluation sample**. This involved deleting a small number of evaluations (four each) of programming in Europe and Central Asia region and Latin America and Caribbean Region and replacing those with evaluations of programming in the Middle East and Northern Africa and Sub-Saharan Africa.

Annex D illustrates the **continued strong fit between the final evaluation sample and the critical parameters of the evaluation universe**, so there is little or no loss in representation along those parameters resulting from the adjustments made to better reflect the distribution of UNICEF expenditures during the review period.

In summary, the sample of 62 evaluations (Annex B)²⁴ represents a modified (and thus purposive), stratified random sample of approximately 30 percent of the evaluation reports

²³ After a quality review- refer Annex E for quality review grid and results.

²⁴ Covers 50 countries

included in the universe. It is meant to be illustrative of the results of UNICEF's efforts to assess development and humanitarian effectiveness during the period 2016 to 2019. Broadly speaking, the goal of the sample is to present a faithful reflection of UNICEF's own findings and conclusions regarding the 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 (presented in Annex A).

The 20 effectiveness sub-criteria described in Table 2, are identical to the criteria used in 2016, DER of UNICEF, which covered evaluations published from 2012 to 2015.²⁵

Supplementary Data

Discussions with UNICEF staff during the inception phase of the review highlighted some subject matter areas which have become a higher priority as a result of Agenda 2030 and the current Strategic Plan. In response, the reviewers will highlight any evaluation findings with reference to:

- Innovation²⁶ covering programming approaches, partnerships (collaborating on new and different solutions), and technology (high- and low-tech) in addition to barriers to innovate and the ability scale.
- Programming targeted toward adolescents and supported by age- and sex-disaggregated data (second decade of life), keeping in mind that UNICEF takes a life-course approach to adolescent development and participation,²⁷ identifying critical risks and opportunities that have implications for the realization of children's rights, from the first decade through the second.
- Climate change mitigation and adaptation covering how UNICEF is responding to climate crisis threatening to roll back progress on child rights and affecting children (including displacement due to extreme weather events, related diseases, and toxic air).
- Addressing the rights and needs of children with disabilities.

Findings in relation to these four areas will not be rated and quantitatively profiled as for the criteria in Table 2 above. Doing so would create a new time series and reduce the comparability of results across the effectiveness review of UNICEF. Rather, the review team will aggregate

²⁵ One criterion (3.3) was added to improve the relevance of the review to UNICEF's programming during the 2016 DER as compared to the 19 sub-criteria used in the 2012 DER (which looked at evaluations from 2008-2011). Although, sub-criteria 6.5 was added to look at how well UNICEF has been using the results of the evaluations to improve humanitarian and development effectiveness, the data collected and analysis was not included in the main text of the final review report. The sub-criterion 6.4 already examines this issue; however, an effort was made to explore whether or not actions were planned by assessing any information on the significance of those actions and how they were followed up.

²⁶ UNICEF has made innovation a corporate priority (Evaluation of Innovation in UNICEF Work, 2019)

²⁷ <https://www.unicef.org/adolescence>

findings reported in relation to the four topics listed and provide a qualitative assessment of lessons learned.

Review Process and Data Analysis

Each evaluation report in the sample were reviewed by a small review team of two reviewers and two senior members (including the team leader). The first task of each reviewer is to assess the quality of the evaluation report under examination using the Quality Review Grid (Annex 1). The reviewers classified the evaluation findings in each evaluation report as they applied to the six evaluation criteria and 20 sub-criteria presented in Annex A. 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 were developed for each evaluation sub-criteria and each level of finding (from highly unsatisfactory to highly satisfactory). The detailed guidelines are presented in Annex J.

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 all reviewed the same evaluation report and discussed and calibrated their findings. There were also regular meetings of the review team to discuss criteria and compare findings during the review process. The team leader and deputy-team leader reviewed completed evaluation review reports during the process to correct any misalignment and convened discussions of the team where there was any question regarding the interpretation of evidence.

During each review, the analysts not only classified evaluation findings, but they also entered into the web-based review grid for each evaluation, the key evidence used to classify evaluation findings. This allowed the senior team members (including the team leader) to assess the accuracy of 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 62 evaluation reports reviewed with qualitative evidence gathered from the review of a sample of thematic evaluations and interviews with UNICEF staff. This approach is entirely consistent with the data gathering and analysis process used in the 2016 and 2013 Reviews and all other completed DERs since 2010.

Limitations

As a meta-synthesis of previously completed evaluation reports, this development effectiveness review is 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 the 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 2016 to 2019. 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:** As noted earlier, 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 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). At most, it may be possible to compare some results across broader aggregations of stratification (for example, results in sub-Saharan Africa versus other regions or between least developed countries (LDCs) and other categories of countries).
- **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 2016 to 2019 will have originated some years before they were evaluated, and publication of the findings may only have occurred sometime after the evaluation fieldwork was completed. This means that the review will need 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 will be able to replicate the 2016 and 2013 process and provide a meaningful overview of the existing evaluative evidence of organizational effectiveness (and how that may have 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 already been 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= 222)		Final Sample Analyzed (n= 62)	
	Number	%	Number	%
2019	10	5%	4	6%
2018	69	31%	21	34%
2017	70	32%	17	27%
2016	73	33%	20	32%

Distribution by Region

Region	Evaluation Population (n= 222)		Final Sample Analyzed (n= 62)	
	Number	%	Number	%
Europe & Central Asia (ECA)	31	14%	6	10%
East Asia & Pacific (EAPR)	23	10%	6	10%
East & Southern Africa (ESAR)	46	21%	16	26%
Latin America & Caribbean (LACR)	32	14%	5	8%
Middle East & North Africa (MENA)	20	9%	10	16%
South Asia (ROSA)	25	11%	5	8%
West & Central Africa (WCAR)	45	20%	14	23%

Distribution by Country Income Classification

Country Income	Evaluation Population (n= 222)		Final Sample Analyzed (n= 62)	
	Number	%	Number	%
High	5	2%	1	2%
Upper-Middle	62	28%	16	26%
Lower-Middle	78	35%	23	37%
Low	75	34%	22	35%
N/A	2	1%	-	-

Distribution by Thematic Area

Thematic Area	Evaluation Population (n= 222)		Final Sample Analyzed (n=62)	
	Number	%	Number	%
Health	64	29%	21	34%
Nutrition	27	12%	7	11%
HIV & AIDS	12	5%	2	3%
Education	80	36%	23	37%
Child Protection	60	27%	15	24%
WASH	41	18%	12	19%
Social Inclusion	67	30%	15	24%
Humanitarian	27	12%	12	19%
Gender	15	7%	5	8%

Note: Only 62% of the evaluations (in the universe) are specific to one thematic area. The other (38%) evaluations cover more than one thematic and/or cross-cutting area. Therefore, the percentages will add up to more than 100%.

Comparison of the Sample with Expenditure by Region

Region	Program Expenses (2018) - \$5.39 billion (% by region)	Final Sample Analyzed – 62 (% by region)
Global	3.6%	-
Europe & Central Asia (ECA)	5.8%	10%
Asia: East Asia & Pacific (EAPR) and South Asia (ROSA)	17.1%	18%
Latin America & Caribbean (LACR)	3.3%	8%
Middle East & North Africa (MENA)	25.1%	16%
Sub-Saharan Africa: West & Central Africa (WCAR) + East and Southern Africa (ESAR)	45.3%	48%

Annex E: Evaluation Quality Review Grid and Results

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

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

The quality review was completed for 63 of the 66 reports selected initially for review. Three of them were not reviewed because they were part of the earlier exclusion criteria – baseline, and mid-term evaluation report. One of them did not have a report. 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.

These three criteria indeed account for 12 of the 40 points available for the overall quality score. Still, for that very reason, the Development Effectiveness Review (DER) approach (used in this Development Effectiveness Review) 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 one evaluation was excluded from the analysis. This approach is consistent with the method used to date in all earlier DERs.

Results of Quality Review Scoring - 2020

Overall evaluation quality scores			Development (Program) 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	33	52.4%	12	6	9.5%
31 - 35	19	30.1%	11	18	28.6%
25 - 30	10	15.9%	10	24	38.1%
21 - 24	(1)	1.6%	9	14	22.5%
16 - 20	-	-	8	(1)	1.6%
11 - 15	-	-	7	-	-
6 - 10	-	-	1-6	-	-
Total	63 (1)	100.0%	Total	63 (1)	100.0%

Out of the total 63 reports reviewed for quality, one was rejected based on quality scores.

The figures in brackets are the number of reports that failed the quality screening. From the table above, it could be seen that there was one report that had a total score of less than 25 and also failed to meet the minimum score of nine in G, H, and I criteria.

Reports (62 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.

For comparison, the quality tables for 2012 and 2016 reviews are presented below.

Results of Quality Review 2016 (Final Sample 66)

Overall evaluation quality scores			Development (Program) 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)	17.6%	10	22 (1)	30.1%
21 - 24	(4)	5.5%	9	16 (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%
Total	73 (7)	100.0%	Total	73 (7)	100.0%

Results of the Quality Review 2012 (Final Sample 62)

Overall evaluation quality scores			Development (Program) effectiveness key criteria (G, H & I)		
Max points = 40 (Min required = 25)	No. of evaluations	% of evaluations	Max points = 13 (Min required = 9)	No. of evaluations	% of evaluations
36 - 40	15	22.7%	13	-	-
31 - 35	36	54.6%	12	6	9.1%
25 - 30	15	22.7%	11	24	36.4%
21 - 24			10	25	37.9%
16 - 20			9	7	10.6%
11 - 15			8	(2)	3.0%
6 - 10	-	-	7	(1)	1.5%
-	-	-	1-6	(1)	1.5%
Total	66	100.0%	Total	66 (4)	100.0%

The figure in parentheses indicate the number of reports that were not reviewed because they did not either meet the overall minimum score of 25 and/or quality assessment criteria minimum score of nine.

The maximum score for G, H & I criteria for the 2020 and 2016 reviews was 12 as compared to 13 for the 2012 review. This was because the separate points for the baseline and comparison groups in the Evaluation Design (H) were combined, based on experience from similar reviews of other UN agencies and consultations with UNICEF to give only one point. However, the total (40) remained the same, as, under Evaluation Criteria (F), the cross-cutting criteria of “equity” was added (for the 2016 and 2020 reviews), and this ensured that the evaluation report could score a maximum of six points for Evaluation Criteria (H) instead of five points as in the 2012 review.

Annex F: Global Evaluations and Corporate Documents Reviewed

A. Global, Regional and Thematic Evaluation Reports

- UNICEF Response to the South Sudan Humanitarian Situation - 2019
- Evaluation Office: Evaluation of Innovation in UNICEF Work - Synthesis Report – 2018
- Independent Panel Review of the UNICEF Response to PSEA - 2018
- Evaluation of UNICEF contribution to health system strengthening - 2018
- Evaluation of UNICEF Girls' Education Portfolio (2009-2015) - 2018
- UNICEF Multi-country Gender-based Violence in Emergencies Program Evaluation – 2016
- Evaluation of the Coverage and Quality of the UNICEF Humanitarian Response in Complex Humanitarian Emergencies - 2018
- Strengthening Child Protection Systems: Evaluation of UNICEF Strategies and Program Performance – 2018
- Influential Evaluations: A Selection of UNICEF Evaluations That Led to Learning and Change - 2018
- Global Evaluation of UNICEF's Drinking Water Supply Programming in Rural Areas and Small Towns, (2006–2016) - 2018
- Participatory review (evaluation) of the UNICEF response to the refugee and migrant crisis in Europe (2015-2017) - 2018
- External Evaluation of UNICEF's Scaling Up Nutrition and Immunization implemented in 13 sub-Saharan African countries over the course of 2013- 2016 - 2017
- TOWARDS IMPROVED EMERGENCY RESPONSES: Synthesis of UNICEF Evaluations of Humanitarian Action (2010 – 2016) - 2017
- Reducing Stunting in Children Under 5 Years of Age: A comprehensive evaluation of UNICEF's strategies and program performance – Global synthesis report - 2017
- Evaluation of UNICEF's response to the Ebola outbreak in West Africa (2014-2015) – 2017
- Evaluation of UNICEF's Response to the Ebola outbreak in West Africa (2014-2015) – 2017
- Evaluation of UNICEF's Response to the Humanitarian Crisis in South Sudan - 2019
- Participatory Review (Evaluation) of the UNICEF Response to the Refugee and Migrant Crisis in Europe (2015 – 2017) - 2018
- Thematic Evaluation of Government interventions and UNICEF contribution to reducing vulnerabilities, strengthening resilience and promoting the rights of adolescent children in Belarus - 2018
- Real Time Evaluation on the Emergency Drought Situation Response in Kenya, (2017) - 2018
- Real-Time Evaluation of UNICEF SCO (Somalia CO) Humanitarian Response to pre-famine Crisis - 2018
- Conducting a Rapid Assessment of MomConnect Project from September 2011 in eThekweni and Umgungundlovu districts in KwaZulu-Natal - 2016

B. Corporate Documents

- Revised Evaluation Policy of UNICEF (2018)
- UNICEF Strategic Plan 2018-2021
- Final Results Framework of the UNICEF Strategic Plan 2018 – 2021
- 2018 – 2021 Strategic Plan – Common Chapter – UNDP, UNFPA, UNICEF, and UN Women
- Theory of Change Paper, UNICEF Strategic Plan 2018 – 2021

- UNICEF Strategic Plan 2014 – 2017
- UNICEF Annual Reports (2015, 2016, 2017 and 2018)

C. UNICEF Reports

- Annual reports of Executive Director on performance and results (for 2016, 2017 and 2018)
- Annual reports on the evaluation function (for 2016, 2017 and 2018)
- Program Division Annual Report (2016, 2017 and 2018)
- Field Results Groups (FRG) Annual Report 2018
- Management perspective on the annual reports on the valuation function (2012, 2013, and 2014)
- An Independent Review of UNICEF Evaluation Report Quality and Selected Trends from 2016 -2018
- UNICEF GEROS Meta-Analysis 2016 and 2017
- UNICEF Gender Action Plan (GAP) 2018-2021
- Annual report on the implementation of the UNICEF GAP 2018-2021
- Annual report on the implementation of the UNICEF GAP 2014 – 2017
- Annual Results Report 2017 – Gender Equity
- Global Annual Results Report 2018 – Humanitarian Action
- Division of Analysis, Planning and Monitoring (DAPM), Office Management Plan 2019-2021
- UNICEF Nutrition Information System Updates – v3 (PowerPoint)
- Report on the implementation of the Quadrennial Comprehensive Policy Review of operational activities for development of the United Nations system – April 29, 2019
- DAC/UNEG Peer review of the evaluation function of UNICEF – July 2017
- UNICEF Evaluation Management response to the Peer review - 2017
- MOPAN 2015-16 Assessments – UNICEF Institutional Report
- UNICEF Management response to MOPAN 2015-16
- UNICEF (2019) – Scaling innovation for every child

Annex G: Key UNICEF Staff Consulted

1. George Laryea-Adjei, Director of Evaluation, Evaluation Office
2. Ada Ocampo, Senior Evaluation Specialist, Evaluation Office
3. Beth Ann Plowman, Senior Evaluation Specialist, Evaluation Office
4. Jane Mwangi, Evaluation Specialist, Evaluation Office
5. Fabio Sabatini, Evaluation Specialist, Evaluation Office
6. Uyen Kim Huynh, Innovation Specialist, Evaluation Office
7. Tami Aritomi, Evaluation Specialist, Evaluation Office
8. Vidya Ganesh, Director, Division, of Data Analysis, Planning and Monitoring (DAPM)
9. Joseph Barnes, Chief Program, Program Monitoring, DAPM
10. Bolton Tarleh Nyema, Deputy Director, Office of Internal Audit and Investigations
11. Yvon Roy, Chief Audit
12. Thomas Dreesen, Education Specialist, Program Division (PD)
13. Megan Tucker, Program Specialist, Children with Disabilities, PD
14. Jumana Haj-Ahmad Chief Adolescent Development and Participation, PD
15. Eri Dwivedi, Child Protection Specialist, PD
16. Chemba Raghavan, Early Childhood Specialist, PD
17. Abheet J. Solomon Senior Program and Planning Adviser, Health Section, PD
18. Tanya Accone, Senior Adviser on Innovation
19. Mathew Varghese, UN Secretariat

Annex H: Comparison of Results by Country Type

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

	HUS	US	S	HS	Total	Valid #	S & HS	Total
LIC	-	10%	60%	30%	100%	20	90%	22
LMIC	-	10%	60%	30%	100%	20	90%	23
UMIC	-	43%	28%	27%	100%	14	53%	16
HIC	-	-	-	100%	100%	1	100%	1

1.2 UNICEF supported development interventions and programs align with national development and humanitarian goals.

	HUS	US	S	HS	Total	Valid #	S & HS	Total
LIC	-	6%	63%	31%	100%	16	94%	22
LMIC	5%	-	45%	50%	100%	20	95%	23
UMIC	-	-	54%	46%	100%	13	100%	16
HIC	-	-	-	100%	100%	1	100%	1

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

	HUS	US	S	HS	Total	Valid #	S & HS	Total
LIC	-	28%	67%	5%	100%	18	72%	22
LMIC	-	9%	68%	23%	100%	22	91%	23
UMIC		8%	84%	8%	100%	13	92%	16
HIC			100%		100%	1	100%	1

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

	HUS	US	S	HS	Total	Valid #	S & HS	Total
LIC	6%	19%	69%	6%	100%	16	75%	22
LMIC	-	19%	67%	14%	100%	21	81%	23
UMIC	-	36%	56%	9%	100%	11	65%	16
HIC	-	-	100%	-	100%	1	100%	1

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

	HUS	US	S	HS	Total	Valid #	S & HS	Total
LIC		5%	85%	10%	100%	21	95%	22
LMIC	-	9%	82%	9%	100%	23	91%	23
UMIC	-	-	100%	-	100%	15	100%	16
HIC	-	-	100%	-	100%	1	100%	1

2.3 UNICEF programs and interventions 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	-	26%	74%	-	100%	19	74%	22
LMIC	-	13%	64%	23%	100%	22	87%	23
UMIC	-	33%	67%	-	100%	12	67%	16
HIC	-	-	100%	-	100%	1	100%	1

2.4 UNICEF activities contributed to significant changes in national development policies and programs (including for disaster preparedness, emergency response and rehabilitation).

	HUS	US	S	HS	Total	Valid #	S & HS	Total
LIC	-	44%	44%	12%	100%	16	56%	22
LMIC	-	17%	75%	8%	100%	14	83%	23
UMIC	-	14%	64%	22%	100%	12	86%	16
HIC	-	-	-	-	-	-	-	1

3.1 Extent UNICEF supported activities effectively address the cross-cutting issue of Gender Equity.

	HUS	US	S	HS	Total	Valid #	S & HS	Total
LIC	-	35%	53%	12%	100%	17	65%	22
LMIC	-	29%	59%	12%	100%	17	71%	23
UMIC	20%	47%	33%	-	100%	15	33%	16
HIC	-	-	100%	-	100%	1	-	1

3.3 Extent UNICEF programming contributes to equity.

	HUS	US	S	HS	Total	Valid #	S & HS	Total
LIC	5%	35%	60%	-	100%	20	60%	22
LMIC	7%	40%	46%	7%	100%	20	53%	23
UMIC	5%	40%	45%	10%	100%	15	55%	16
HIC	-	-	100%	-	100%	1	-	1

4.1 Benefits continuing or likely to continue after intervention or program completion or there are effective measures to link the humanitarian relief operations, to rehabilitation.

	HUS	US	S	HS	Total	Valid #	S & HS	Total
LIC	15%	60%	20%	5%	100%	20	25%	22
LMIC	10%	52%	19%	19%	100%	21	38%	23
UMIC	13%	54%	33%	-	100%	15	33%	16
HIC	-	100%	-	-	100%	1	-	1

4.2 Extent UNICEF supported interventions and programs are reported as sustainable in terms of institutional and/or community capacity.

	HUS	US	S	HS	Total	Valid #	S & HS	Total
LIC	17%	11%	67%	5%	100%	18	72%	22
LMIC	-	18%	77%	5%	100%	22	82%	23

UMIC	-	40%	47%	13%	100%	15	60%	16
HIC	-	-	100%	-	100%	1	100%	1

4.3 Extent UNICEF development programming contributes to strengthening the enabling environment for development.

	HUS	US	S	HS	Total	Valid #	S & HS	Total
LIC	7%	33%	53%	7%	100%	15	60%	22
LMIC	-	30%	65%	5%	100%	20	70%	23
UMIC	-	33%	59%	8%	100%	12	67%	16
HIC	-	100%	-	-	100%	1	-	1

5.1 Program activities are evaluated as cost/resource-efficient.

	HUS	US	S	HS	Total	Valid #	S & HS	Total
LIC	11%	47%	21%	21%	100%	19	42%	22
LMIC	-	40%	50%	10%	100%	20	60%	23
UMIC	8%	61%	23%	8%	100%	13	31%	16
HIC	-	-	100%	-	100%	1	100%	1

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
LIC	6%	56%	32%	6%	100%	16	40%	22
LMIC	-	44%	45%	11%	100%	18	56%	23
UMIC	28%	29%	43%	-	100%	7	43%	16
HIC	-	-	-	-	-	-	-	1

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).

	HUS	US	S	HS	Total	Valid #	S & HS	22
LIC	28%	39%	33%	-	100%	18	33%	23
LMIC	12%	31%	44%	13%	100%	16	57%	16
UMIC	-	75%	25%	-	100%	8	25%	1
HIC	-	-	100%	-	100%	1	100%	1

6.1 Systems and processes for evaluation are effective.

	HUS	US	S	HS	Total	Valid #	S & HS	22
LIC	-	29%	59%	12%	100%	17	71%	23
LMIC	-	14%	62%	24%	100%	21	86%	16
UMIC	15%	15%	62%	8%	100%	13	70%	1
HIC	-	-	100%	-	100%	1	100%	1

6.2 Systems and processes for monitoring and reporting on program results are effective.

	HUS	US	S	HS	Total	Valid #	S & HS	22
LIC	-	65%	35%	-	100%	20	35%	23
LMIC	-	45%	46%	9%	100%	22	55%	16

UMIC	12%	63%	25%	-	100%	16	25%	1
HIC	-	-	-	100%	100%	1	100%	1

6.3 Results-Based Management (RBM) systems are effective.

	HUS	US	S	HS	Total	Valid #	S & HS	22
LIC	5%	74%	21%	-	100%	19	21%	23
LMIC		55%	27%	18%	100%	22	45%	16
UMIC	25%	50%	25%	-	100%	16	25%	1
HIC	-	-	-	100%	100%	1	100%	1

6.4 UNICEF makes use of evaluation to improve development/humanitarian effectiveness.

	HUS	US	S	HS	Total	Valid #	S & HS	Total
LIC	-	4%	64%	32%	100%	22	96%	22
LMIC	-	-	43%	57%	100%	23	100%	23
UMIC	-	6%	38%	56%	100%	16	94%	16
HIC	-	-	-	100%	100%	1	100%	1

Annex I: Guide for Review Team to Classify Evaluation Findings

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
Common Development and Humanitarian Evaluation Assessment Criteria				
1. Relevance				
1.1 UNICEF supported programs, interventions and initiatives are suited to the needs and/or priorities of the target group	Evaluation finds that substantial elements of program or intervention and/or its 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 intervention activities and outputs and the needs and priorities of the target group.	Evaluation finds that the UNICEF supported program or intervention 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 humanitarian operations) and that the resulting outputs are designed to meet the needs of the target group.	Evaluation finds that credible methods (including needs assessment for relief operations) were used to identify target group needs and priorities (including consultations with target group members). Further it finds that the program or intervention takes is designed to meet those needs and priorities (whether or not it does so successfully).
1.2 UNICEF supported humanitarian and development programs and interventions align with national humanitarian and development goals:	The evaluation reports that significant elements of UNICEF supported humanitarian and development programs or interventions run counter to national humanitarian and development priorities with a resulting loss of effectiveness.	The evaluation reports a <u>significant portion</u> (1/4 or more) of UNICEF supported humanitarian and development programs and subject to the evaluation are not aligned with national plans and priorities but there is no evidence that they run counter to those priorities.	<u>Most</u> UNICEF supported humanitarian and development programs and interventions are reported in the evaluation to be fully aligned with national plans and priorities as expressed in national development and sector plans and priorities. If some programs and interventions are reported in the evaluation as not directly supportive of national plans and priorities, they do not run counter to those priorities.	<u>All</u> evaluated humanitarian and development interventions 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.
1.3 UNICEF has developed effective partnerships with governments, bilateral and multilateral development and	Evaluation finds that UNICEF experiences significant divergence in priorities from those of its (government, non-	Evaluation finds that UNICEF has experienced significant difficulties in developing effective relationships with	Evaluation finds that UNICEF has improved the effectiveness of its relationship with partners over time during the evaluation	Evaluation finds that UNICEF has consistently achieved a high level of partnership during the evaluation period.

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
humanitarian organizations and non-governmental organizations for planning, coordination and implementation of support to development and/or emergency preparedness, and humanitarian relief and rehabilitation efforts.	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.	partners and that there has been significant divergence in the priorities of UNICEF and its partners.	period and that this partnership was effective at the time of the evaluation or was demonstrably improved.	
2. Achievement of Humanitarian and Development Objectives and Expected Results				
2.1 UNICEF supported programs and interventions 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 interventions either achieve at least a majority of stated output and outcome objectives (more than 50% if stated) or the most important of stated output and outcome objectives are achieved.	UNICEF supported programs and interventions achieve all or almost all significant development and/or humanitarian objectives at the output and outcome level. (If impact level results are found please note)
2.2 UNICEF supported programs and interventions 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 interventions and programs result in 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 humanitarian situation.	UNICEF supported interventions 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 humanitarian situation.	UNICEF supported interventions 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 humanitarian situation.
2.3 UNICEF programs and interventions made differences for a substantial number of beneficiaries and where	UNICEF supported interventions and programs have not contributed to positive changes in the lives of	UNICEF supported interventions and programs have contributed to positive changes in the lives of only a small number of beneficiaries	UNICEF supported interventions and programs have contributed to positive changes in the lives of substantial numbers of	UNICEF supported interventions and programs have contributed to positive changes in the lives of substantial numbers of

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
appropriate contributed to national development goals.	beneficiaries as measured quantitatively or qualitatively.	(when compared to intervention or program targets and local or national goals if established).	beneficiaries as measured quantitatively or qualitatively. These may result from development or humanitarian operations and may include the avoidance of negative effects of emergencies.	beneficiaries. Further, they have contributed to the achievement of specific national development goals or have contributed to meeting humanitarian relief objectives agreed to with the national government and/or national and international development and humanitarian organizations.
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 reforms.	National policies and programs in a given sector or area of development (including disaster preparedness, and humanitarian action) were deficient and required strengthening but UNICEF activities have not addressed these deficiencies.	UNICEF activities have not made a significant contribution to the development of national policies and programs in a given sector or area of development or humanitarian action. (Policy changes in humanitarian situations may include allowing access to the affected populations).	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 rehabilitation.	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 rehabilitation. Further, the supported policies and program implementation modalities have resulted in improved positive impacts for target group members.
3. Cross Cutting Themes: Inclusive Humanitarian Assistance and Development which can be Sustained				
3.1 Extent UNICEF supported programs, interventions and activities effectively address the cross-cutting issue of Gender Equity.	UNICEF supported programs and interventions are unlikely to contribute to Gender Equity or may in fact contribute to increases in gender inequalities.	UNICEF supported activities either lack Gender Equity objectives or achieve less than half of their stated Gender Equity 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).	UNICEF supported programs and interventions achieve a majority (more than 50%) of their stated Gender Equity objectives.	UNICEF supported programs and interventions achieve all or nearly all of their stated Gender Equity objectives.
3.2 Extent changes are environmentally sustainable.	UNICEF supported programs and interventions do not	UNICEF supported programs and interventions do not	UNICEF supported programs and interventions include some	UNICEF supported programs and interventions are

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
	include planned activities or intervention design criteria intended to promote environmental sustainability. In addition, changes resulting from UNICEF supported programs and interventions are not environmentally sustainable.	include planned activities or intervention design criteria intended to promote environmental sustainability. There is, however, no direct indication that intervention or program results are not environmentally sustainable. OR UNICEF supported programs and interventions include planned activities or intervention design criteria intended to promote sustainability but these have not been successful.	planned activities and intervention design criteria to ensure environmental sustainability. These activities are implemented successfully and the results are environmentally sustainable. <i>Special note should be taken of any efforts to address climate change in the supported program or intervention – including adaptation as well as mitigation efforts.</i>	specifically designed to be environmentally sustainable and include substantial planned activities and intervention design criteria to ensure environmental sustainability. These plans are implemented successfully and the results are environmentally sustainable. S <i>Special note should be taken of program or intervention features aimed to support countries and communities in adapting to climate change or mitigating CO2 emissions.</i>
3.3. The extent to which UNICEF supported programs effectively address issues of equity for children and adolescents (in either or both of the first two decades of life).	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, children with disability, or urban/rural divide. Further, the program did not differentiate between barriers to participation facing children and adolescents (first or second decade of life). 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.	The evaluation reports that the program identified one or more barriers to equity based on gender, poverty, ethnicity, children with disability, etc. but the program elements designed to help marginalized or under-served groups overcome these barriers have been shown to be inadequate or have not resulted in any improvement in access.	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. <i>Special attention should be paid to efforts to reach marginalized and disadvantaged adolescents (aged 11 to 19) and children with disabilities.</i>	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.
4. Sustainability				

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
4.1 Benefits are continuing or likely to continue after intervention or program completion or there are effective measures to link the humanitarian relief operations, to rehabilitation, reconstructions and, eventually, to longer-term developmental results.	There is a very low probability that the program/intervention will result in continued intended benefits for the target group after intervention completion. For humanitarian operations, the evaluation finds no strategic or operational measures to link relief, to rehabilitation, reconstruction and, eventually, to development.	There is a low probability that the program/intervention will result in continued benefits for the target group after completion. For humanitarian relief operations, efforts to link the relief phase to rehabilitation, reconstruction and, eventually, to development are 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 should be given)	It is likely that the program or intervention will result in continued benefits for the target group after completion. For humanitarian relief operations, the strategic and operational measures to link relief to rehabilitation, reconstruction and, eventually, development are credible.	It is highly likely that the program or intervention will result in continued benefits for the target group after completion. For humanitarian relief operations, the strategic and operational measures to link relief to rehabilitation, reconstruction and, eventually, development are credible. Further, they are likely to succeed in securing continuing benefits for target group members.
4.2 Extent UNICEF supported interventions and programs are reported as sustainable in terms of institutional and/or community capacity.	The design of UNICEF supported programs and interventions 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 interventions failed to take account of identified needs to strengthen local capacities for delivery of relief operations and/or for managing the transition to rehabilitation and/or development.	UNICEF programs and interventions may have failed to contribute to strengthening institutional and/or community capacity (including service delivery capacity) or, in the case of humanitarian operations, were not appropriate, to strengthen national or local capacities for delivery of relief operations and/or for managing the transition to rehabilitation and/or development.	UNICEF programs and interventions have contributed to strengthening institutional and/or community capacity but with limited success.	Either UNICEF programs and interventions 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.
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	UNICEF development activities and/or UNICEF supported interventions and programs have not made a notable contribution to changes in the enabling environment for development.	UNICEF development activities and/or UNICEF supported interventions 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 planning to	UNICEF development activities and/or UNICEF supported interventions 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

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
	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 interventions failed to address the identified weakness successfully, further limiting program results.		meet the needs of children and adolescents; the legal framework for child protection, 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.	planning for children and adolescents; the legal framework for child protection, systems and processes for public consultation and for participation by civil society in planning for meeting the needs of children and adolescents; 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.
5. Efficiency				
5.1 Program activities are evaluated as cost/resource efficient (including provision of value for money):	Credible information indicating that UNICEF supported programs and interventions (development, emergency preparedness, relief and rehabilitation) are not cost/resource efficient and/or do not provide sufficient value for money in terms of results achieved in relation to expenditures.	UNICEF supported programs and interventions under evaluation (development and humanitarian) 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 programs and interventions under evaluation present mixed findings on the cost/resource efficiency of the inputs.	Level of outputs and outcomes achieved (development, and humanitarian) 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 and humanitarian programs and interventions are designed to include activities and inputs that produce outputs in the most cost/resource efficient manner available at the time. Or, there is a credible assertion in the evaluation that the program or intervention represents significantly positive value for money.
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 interventions are achieved on time, there is no credible	Less than half of stated output and outcome level objectives of UNICEF supported programs and interventions are achieved on time but the program or	More than half of stated output and outcome level objectives of UNICEF supported programs and interventions are achieved on time and this level is	Nearly all stated output and outcome level objectives of UNICEF supported programs and interventions are achieved on time or, in the case of

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
	plan or legitimate explanation found by the evaluation which would suggest significant improvement in on-time objectives achievement in the future.	intervention 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.	appropriate to the context faced by the program during implementation, particularly for humanitarian programming.	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 intervention/program implementation and follow up are efficient (including systems for engaging staff, procuring intervention inputs, disbursing payment, logistical arrangements etc.)	Serious deficiencies in agency systems and procedures for project/program implementation that result in significant delays in intervention start-up, implementation or completion and/or significant cost increases.	Some deficiencies in agency systems and procedures for project/program implementation but no indication that these have contributed to delays in achieving project/program objectives.	Agency systems and procedures for intervention implementation are reasonably efficient and have not resulted in significant delays or increased costs.	The efficiency of agency systems and procedures for intervention implementation represent an important organizational strength in the implementation of the program under evaluation.
6. Using Evaluation and Monitoring to Improve humanitarian and development Effectiveness				
6.1 Systems and process for evaluation are effective.	Evaluation practices in use for programs and interventions of this type (development and humanitarian) are seriously deficient.	No indication that programs and interventions of this type (development and humanitarian) 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 humanitarian programming.	Program being evaluated (along with similar programs and projects) is subject to systematic regular evaluations or describes significant elements of such practice.
6.2 Systems and processes for monitoring and reporting are effective	Absence of monitoring and reporting systems for development and humanitarian programs and interventions under evaluation. This would include the absence of adequate monitoring of outputs during	While monitoring and reporting systems for 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.

Criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
	the implementation of humanitarian programming.			
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 J: Changes Over Time Weighted Average Scores 2012, 2016 and 2020

Stable Results Over Time with Some Variations

The general pattern of positive and negative findings for most effectiveness criteria has remained quite stable across the ten years of UNICEF evaluations reports published from 2009 to 2019 and reviewed in 2012, 2016 and 2019. However, for a few sub-criteria, there have been either steady improvement or significant shifts in results. Table below provides another perspective on these changes by noting the portion of satisfactory and highly satisfactory findings for selected criteria in each of the three reviews. This is accompanied by the weighted average score for each sub-criterion as described in Section 4 in Box 10.

Selected Sub-Criteria: Significant Changes Over Time	Positive Percent 2012	Weighted Average Score 2012	Positive Percent 2016	Weighted Average Score 2016	Positive Percent 2020	Weighted Average Score 2020
1.3 Effective Partnerships	87 %	3.10	70 %	2.85	85%	2.98
3.1 Effectively Addressing Gender Equity	47 %	2.38	48 %	2.43	58 %	2.60
4.1 Sustainability of Benefits	51 %	2.46	36 %	2.25	32 %	2.29
6.1 Effective Systems for Evaluation	64 %	2.78	70 %	2.79	77 %	2.88

For two of the sub-criteria in table three (3.1 on Gender Equity and 6.1 on evaluation systems) there is a pattern of steadily increasing frequency of positive evaluation findings reported across the three effectiveness reviews. For sub-criterion 1.3 on effective partnerships the pattern is one of decline from 2012 to 2016 but a significant rebound in 2020. Finally, for sub-criterion 4.1 on the sustainability of program benefits, the pattern is one of steady decline from just over half of evaluations reporting positive results in 2012 to 32 percent in 2020.

Criteria and Sub-Criteria	Weighted Average Scores		
	2102	2016	2020
Relevance			

1.1 Programs suited to the needs and priorities of the target group.	3.17	2.90	3.13
1.2 Programs align with national development goals.	3.44	3.51	3.38
1.3 Effective partnerships with governments, bilateral and multilateral development and humanitarian organizations and NGOs.	3.10	2.85	2.98
Objectives Achievement			
2.1 Programs and activities achieve their stated humanitarian and development objectives.	2.86	2.88	2.81
2.2 Programs and activities have resulted in positive benefits for target group members.	3.04	3.08	3.02
2.3 Programs and activities made differences for a substantial number of beneficiaries.	2.98	2.84	2.87
2.4 Programs contributed to significant changes in national humanitarian and development policies and programs.	3.04	3.02	2.88
Cross-Cutting Themes			
3.1 Extent UNICEF support effectively addresses Gender Equity.	2.38	2.43	2.60
3.2 Extent to which changes are environmentally sustainable.	N/A	N/A	N/A
3.3 Programs effectively address issues of equity.	N/A	2.65	2.57
Sustainability			
4.1 Benefits likely to continue after project or completion.	2.46	2.25	2.29
4.2 Programs and activities are reported as sustainable for institutional and/or community capacity.	2.45	2.69	2.73
4.3 Programming contributes to strengthening the enabling environment for development.	2.71	2.69	2.66
Efficiency			
5.1 activities are evaluated as cost/resource efficient.	2.70	2.78	2.54
5.2 Implementation and objectives achieved on time	2.38	2.67	2.44
5.3 Systems and procedures for implementation and follow up are efficient.	2.23	2.38	2.31
Using Monitoring and Evaluation to Improve Effectiveness			
6.1 Systems and process for evaluation are effective.	2.78	2.79	2.88
6.2 Systems and processes for monitoring and reporting are effective.	2.39	2.36	2.43
6.3 Results based management systems are effective.	N/A	2.28	2.35
6.4 Evaluation is used to improve humanitarian and development effectiveness.	2.55	3.21	3.42

