

EVALUATION REPORT

UNICEF STRATEGIC PLAN, 2014-2017 EVALUABILITY ASSESSMENT

Final Report

EVALUATION OFFICE
July 2015



EVALUATION REPORT

**UNICEF STRATEGIC
PLAN, 2014-2017
EVALUABILITY
ASSESSMENT**

Final Report

EVALUATION OFFICE
July 2015

unite for
children



UNICEF Strategic Plan, 2014-2017 Evaluability Assessment: Final Report

© United Nations Children's Fund (UNICEF), New York, 2015

United Nations Children's Fund

Three United Nations Plaza

New York, New York 10017

July 2015

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

This assessment was prepared by independent consultants Beth Plowman and Philip Setel. Mathew Varghese, Senior Evaluation Officer, Evaluation Office, led and managed the overall assessment process with substantial support from Laurence Reichel and Laura Olsen and in close collaboration with the Division of Data, Research and Policy.

The purpose of the report is to facilitate the exchange of knowledge among UNICEF personnel. The contents of the report do not necessarily reflect the policies or views of UNICEF. The designations in this publication do not imply an opinion on the legal status of any country or territory, or of its authorities, or the delimitation of frontiers.

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

For further information, please contact:

Evaluation Office
United Nations Children's Fund
Three United Nations Plaza
New York, New York 10017
evalhelp@unicef.org

Preface

UNICEF is mandated by the United Nations General Assembly to advocate for the protection of children's rights, to help meet their basic needs and to expand their opportunities to reach their full potential. The Executive Board approved the UNICEF Strategic Plan, 2014-2017, *Realizing the Rights of Every Child, Especially the Most Disadvantaged*, in September 2013. It is a critically important document that guides UNICEF-supported programmes in all countries and for work plans and office management plans in all regional offices and Headquarters divisions.

The UNICEF Evaluation Office commissioned this evaluability assessment to determine the extent to which progress towards the objectives set out in the Strategic Plan can be readily and reliably measured, monitored and evaluated. Evaluability assessments can provide assurance to UNICEF management and stakeholders that a plan is robust and that it will be possible to measure the results achieved following implementation. This exercise was conducted early enough in the programme cycle to allow UNICEF to strengthen, where necessary, the systems needed to measure, monitor and evaluate progress for results.

The assessment has generated interesting findings that I believe will strengthen the Strategic Plan and, ultimately, UNICEF's organizational and development effectiveness. It also raises questions for the organization to consider in the following programme cycle and will prove useful for future evaluations of all UNICEF programming.

The Evaluation Office commissioned two independent consultants, Beth Plowman and Philip Setel, to undertake the assessment. On behalf of the Evaluation Office I would like to thank both consultants for their excellent work. I would also like to thank the Evaluation Office colleagues who managed the assessment: Mathew Varghese, Laurence Reichel and Laura Olsen. Finally, the assessment was conducted in a participatory manner and I am grateful to UNICEF staff at all levels of the organization for their interest and support throughout the process, especially colleagues in the Division of Data, Research and Policy.

Colin M. Kirk
Director
Evaluation Office
UNICEF New York Headquarters

Contents

Preface

Contents

Acronyms

Glossary

Executive summary

Resumen Ejecutivo

Resume Analytique

1. Introduction and background

1.1. The UNICEF Strategic Plan, Theory of Change and results framework

2. Approach, objectives and methods

2.1. Approach of the evaluability assessment

2.2. Objectives

2.3. Methods

2.4. Limitations of the evaluability assessment

3. Evaluability of the UNICEF Strategic Plan in principle

3.1. Relevance

3.2. The Strategic Plan as a framework

3.3. Clarity of the causal chain

3.4. Key assumptions, risks and mitigation strategies

4. Evaluability of the UNICEF Strategic Plan in practice

4.1. Indicators, baselines and targets

4.2. Quality of the indicators and implications

4.3. Evaluability of normative principles and crosscutting issues

4.4. Data generating and reporting systems

4.5. Systems to link resources to results

5. Translating the Strategic Plan to regional and country office levels

6. Suggestions for future evaluations

7. Implications for UNICEF

Annexes

Annex 1 Evaluability assessment inception report

Annex 2 Document review/extraction templates

Annex 3 Documents reviewed

Annex 4 Individuals interviewed

- Annex 5 Strategic Plan characteristics
- Annex 6 Questions on Theory of Change evidence base
- Annex 7 Summary of assumptions by Strategic Plan outcome
- Annex 8 Indicator analyses
- Annex 9 Regional office priorities (as per Regional Office Management Plans)

List of tables

- Table 1: Evaluability assessment objectives and key questions
- Table 2: Criteria to assess evaluability in principle
- Table 3: Evaluability in practice: Dimensions of indicator quality
- Table 4: Relevance of the Strategic Plan to existing commitments and the Sustainable Development Goals
- Table 5: Outputs grouped by varied bases
- Table 6: Gaps in the causal chain for specific outcomes
- Table 7: Gaps in the causal chain common across outcomes
- Table 8: Assessment of baseline in place
- Table 9: Assessment of target identified
- Table 10: Assessment of indicator quality: Is indicator defined?
- Table 11: Assessment of indicator quality: Can indicator be reliably measured?
- Table 12: Assessment of indicator quality: Is indicator specific?
- Table 13: Assessment of indicator quality: Is indicator sensitive to change?
- Table 14: Data quality checks
- Table 15: Attributes used to code country programme outputs
- Table 16: Attributes used to code country programme activities

List of figures

- Figure 1: Theory of Change logical structure
- Figure 2: Criteria-based assessment of indicator quality, 86 indicators in results framework
- Figure 3: Structure of country reporting system on performance indicators

Acronyms

CCCs	Core Commitments for Children in Humanitarian Action
CEE/CIS	Central and Eastern Europe and the Commonwealth of Independent States
COAR	Country Office Annual Report
CPD	country programme document
CPAP	country programme action plan
CRC	Convention on the Rights of the Child
CRING	Country Reporting on Indicators for the Goals
EAPRO	East Asia and the Pacific Regional Office
ESARO	Eastern and Southern Africa Regional Office
LACRO	Latin America and the Caribbean Regional Office
MENARO	Middle East and North Africa Regional Office
MoRES	Monitoring Results for Equity System
MTR	mid-term review of the Strategic Plan 2014-2017
MTSP	medium-term strategic plan, 2006-2013
QCPR	Quadrennial Comprehensive Policy Review
RAM	Results Assessment Module
ROSA	Regional Office for South Asia
SDG	Sustainable Development Goal
SMQ	strategic monitoring questions
UBRAF	Unified Budget, Results and Accountability Framework
UNAIDS	Joint United Nations Programme on HIV/AIDS
UNDAF	United Nations Development Assistance Framework
UNDG	United Nations Development Group
UNDP	United Nations Development Programme
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNFPA	United Nations Population Fund
UN Women	United Nations Entity for Gender Equality and the Empowerment of Women
WASH	Water supply, sanitation and hygiene
WCARO	West and Central Africa Regional Office
WFP	World Food Programme
WHO	World Health Organization

Glossary

Term	Definition	Source
Assumptions	<p>Hypotheses about factors or risks that could affect the progress or success of a development intervention.</p> <p>Note: Assumptions can also be understood as hypothesized conditions that bear on the validity of the evaluation itself, e.g., about the characteristics of the population when designing a sampling procedure for a survey. Assumptions are made explicit in theory-based evaluations where evaluation systematically tracks the anticipated results chain.</p>	OECD, Glossary of Key Terms in Evaluation and Results Based Management
Benchmark	<p>Reference point or standard against which performance or achievements can be assessed.</p> <p>Note: A benchmark refers to the performance that has been achieved in the recent past by other comparable organizations, or what can be reasonably inferred to have been achieved under the circumstances.</p>	OECD, Glossary of Key Terms in Evaluation and Results Based Management
Evaluability	<ol style="list-style-type: none"> 1. During the planning stage of an undertaking, evaluation functions can contribute to the process by improving the ability to evaluate the undertaking and by building an evaluation approach into the plan. To safeguard independence this should be performed in an advisory capacity only. 2. Before undertaking a major evaluation requiring a significant investment of resources, it may be useful to conduct an evaluability exercise. This would consist of verifying if there is clarity in the intent of the subject to be evaluated, sufficient measurable indicators. 	UNEG, Norms for Evaluation in the UN System
Evaluation	<p>An evaluation is an assessment, as systematic and impartial as possible, of an activity, project, programme, strategy, policy, topic, theme, sector, operational area, institutional performance, etc. It focuses on expected and achieved accomplishments, examining the results chain, processes, contextual factors and causality, in order to understand achievements or the lack thereof. It aims to determine the relevance, impact, effectiveness, efficiency and sustainability of the interventions and contributions of the organizations of the United Nations system. An evaluation should provide evidence-based information that is credible, reliable and useful, enabling the timely incorporation of findings, recommendations and lessons into the decision-making processes of the organizations of the United Nations system and its members.</p>	UNEG, Norms for Evaluation in the UN System
Impacts	<p>Positive and negative, primary and secondary, long-term effects produced by a development intervention, directly or indirectly, intended or unintended.</p>	OECD, Glossary of Key Terms in Evaluation and Results Based Management

Indicator	Quantitative or qualitative factor or variable that provides a simple and reliable means to measure achievement, to reflect the changes connected to an intervention, or to help assess the performance of a development actor.	OECD, Glossary of Key Terms in Evaluation and Results Based Management
Monitoring	Management's continuous examination of progress achieved during the implementation of an undertaking to track compliance with the plan and to take necessary decisions to improve performance.	UNEG, Norms for Evaluation in the UN System
Outcome	The likely or achieved short-term and medium-term effects of an intervention's outputs. Related terms: result, outputs, impacts, effect.	UNEG, Norms for Evaluation in the UN System
Output	The products, capital goods and services that result from a development intervention; may also include changes resulting from the intervention that are relevant to the achievement of outcomes.	OECD, Glossary of Key Terms in Evaluation and Results Based Management
Performance measurement	A system for assessing the performance of development interventions against stated goals. Related terms: performance monitoring, indicator.	OECD, Glossary of Key Terms in Evaluation and Results Based Management
Results	The output, outcome or impact (intended or unintended, positive and/or negative) of a development intervention. Related terms: outcome, effect, impact.	OECD, Glossary of Key Terms in Evaluation and Results Based Management
Results-based management	A management strategy focusing on performance and achievement of outputs, outcomes and impacts. Related term: logical framework.	OECD Glossary of Key Terms in Evaluation and Results Based Management
Results framework	The programme logic that explains how the development objective is to be achieved, including causal relationships and underlying assumptions. Related terms: results chain, logical framework.	OECD, Glossary of Key Terms in Evaluation and Results Based Management
Results chain	The causal sequence for a development intervention that stipulates the necessary sequence to achieve desired objectives beginning with inputs, moving through activities and outputs, and culminating in outcomes, impacts and feedback. In some agencies, reach is part of the results chain. Related terms: assumptions, results framework.	OECD, Glossary of Key Terms in Evaluation and Results Based Management
Risk analysis	An analysis or an assessment of factors (called assumptions in the log frame) that affect or are likely to affect the successful achievement of an intervention's objectives. A detailed examination of the potential unwanted and negative consequences to human life, health, property or the environment posed by development interventions; a systematic process to provide information regarding such undesirable consequences; the process of quantification of the probabilities and expected impacts for identified risks.	OECD, Glossary of Key Terms in Evaluation and Results Based Management
Theory of Change or	What are the underlying rationales and assumptions or theory that defines the relationships or chain of results	UNDP, Handbook on

results/outcome map	that lead initiative strategies to intended outcomes? What are the assumptions, factors or risks inherent in the design that may influence whether the initiative succeeds or fails?	Planning, Monitoring and Evaluating for Development Results.
----------------------------	--	--

Executive summary

The UNICEF Evaluation Office has commissioned an evaluability assessment of the UNICEF Strategic Plan, 2014-2017 at an early stage of the Strategic Plan's implementation. Broadly stated, an evaluability assessment is a review of a proposed activity to determine whether its objectives are adequately defined and its results are verifiable. The evaluability assessment provides UNICEF management and stakeholders with evidence regarding how robust the plan is, the degree to which indicators are validated and measurable and the extent to which systems are in place to measure and verify performance. This document presents the findings and recommendations of the evaluability assessment, which was conducted from September 2014 to February 2015. It is appreciated that enormous effort was required to create the Strategic Plan and its associated documents. It is hoped that the evaluability assessment can add value in by improving this impressive set of materials.

The evaluability assessment was framed around the following five objectives:

1. Assess the relevance, clarity and coherence of the Strategic Plan's design, including the integrity of the causal chain linking inputs and activities with expected outputs, outcomes and impact.
2. Assess the adequacy and validity of the indicators, tools and systems for monitoring, measuring and verifying results; and the use of systems and tools to assess the adequacy and quality of data to be used for monitoring, measuring and evaluating results, including significant gaps in coverage.
3. Assess data and systems to determine the degree to which human and financial resources can be evaluated against the expected results.
4. Provide guidance on approaches to the evaluation of the Strategic Plan with a view to enhancing its utility.
5. Provide evidence and recommendations to guide any enhancements that may be required regarding the design of the Strategic Plan.

The evaluability assessment was guided by an analytical construct centred on 'evaluability in principle' based on the internal logic of a programme, as represented by its theory of change, and 'evaluability in practice', as represented by the measurement approach, tools and indicators used to reflect performance and results achieved. The evaluability assessment relied on a mix of methods and tools to boost rigour and standardization while also integrating qualitative information to contextualize and nuance the findings. Based on best practices identified in the literature, the evaluability assessment made use of rubrics as tools for review of core strategy documents. Semi-structured interviews with a range of UNICEF staff members at Headquarters and regional and country offices provided additional critical information. Two regional offices were visited to answer questions on the regional programme and the translation of the Strategic Plan into country programmes. Finally, several questions required the examination of internal systems for results reporting and drew upon discussion of systems elements with staff and the development of source maps to identify the directional flow of results data from source to Headquarters.

Ten key findings emerged from the evaluability assessment. These key findings are

summarized below along with implications and recommendations:

1. The Strategic Plan is widely viewed as the most coherent and technically sound that UNICEF has developed to date.
2. The Strategic Plan is relevant to the organizational mandate and international agreements to which UNICEF is bound, and is aligned with the Sustainable Development Goals (SDGs).
3. The Strategic Plan serves as an overarching framework.
4. A significant challenge to evaluability lies in the lack of explicitly stated logical connections of the causal chain from outputs to outcomes and from outcomes to impact. While results statements appear at each level of the Theory of Change, there is very little description of how one level leads to or contributes to another. In other words, while the structure of a results hierarchy is present, the 'if-then' statements that are needed to plausibly link one level to another are largely absent.
5. The indicators examined generally performed well against criteria used for this exercise, including indicator definition, existence of baseline and specificity to the intended result. However, a few performance indicators found in the results framework are weaker and, in some cases, may impede evaluation through poor quality or limited interpretability.
6. The entire Strategic Plan is framed with an equity lens, yet the approach to measurement of equity is weak in the Strategic Plan and is further hampered by the questionable availability of data to measure progress in this area.
7. Crosscutting themes pose challenges for evaluability.
8. Data sources and systems for indicator measurement are identified and were either in place or in the process of being established at the time of the evaluability assessment. Many lessons learned have been taken on board as these indicator measurement systems were rolled out and important steps have been taken to further refine and improve them. Due to timing, the evaluability assessment was not able to examine a full cycle of reporting that included these revisions. Some areas of concern remain regarding the responses to the strategic monitoring questions (SMQ) serving as the basis for the majority of performance indicator reporting.
9. Although current reporting systems categorize financial commitments by country office outputs, programme areas, key activities, sub-activities and implementation strategies, the materials reviewed did not specify how resource flows, across levels, would be tracked according to the results of the Strategic Plan (i.e. outcomes and outputs). Importantly, the resource requirement to achieve the anticipated results is not articulated.
10. Country offices, with the strong support of regional offices, have already utilized a number of opportunities in the country programme cycle to reflect on the Strategic Plan and make adjustments in their programmes. Although not outlined in the Strategic Plan, regional offices have facilitated the uptake of strategic plan elements in country programmes by interpreting the global Strategic Plan in terms of regionally-specific priorities.

The analytical lens used to assess the current Strategic Plan examined elements in place

taking into account relevance and plausibility, as well as structural and contextual issues. A strategic plan is a product of a planning process that defines an organization's medium- to long-term goals, as well as the resources required to achieve them. A strategic plan serves to establish and implement a vision within an organization and to support the organization in making choices and responding to emerging issues, needs and mandates. A strategic plan should contain the following elements: a) clear objectives; b) direction and focus; c) well-defined principles or normative issues, as well as standards for measuring them; d) articulation of the implementation strategies; e) alignment of capacity and resources to match with objectives, focus and implementation strategies; f) explicit identification of results and indicators for measuring progress.

The evaluability assessment finds that the objectives and direction of the Strategic Plan are clear. It notes that the standards for measuring normative issues are weak. The evaluability assessment highlights the lack of specifically stated linkages between outputs and outcomes and between outcomes and impact. This lack of clarity about how each outcome links to the organization's intended impact means that the Strategic Plan cannot be used to make trade-offs or weigh decisions across the outcome areas should the question or necessity arise. The relationship between outputs and outcomes is equally problematic as the lack of specifically identified linkages means it is not possible to establish how outputs contribute to outcomes.

The evaluability assessment notes that UNICEF cannot currently determine the adequacy of resources to achieve the anticipated results. This challenge lies not in the reporting systems but follows from the use of a budgeting approach that is means-based (based on revenues supplied) rather than results-based (providing an aggregate costing of work necessary to achieve results). This places the organization at some risk, as the results to be achieved are not associated with evidence-based resource requirements.

The evaluability assessment finds that the Strategic Plan serves as an overarching framework. The analysis notes that while the Strategic Plan looks like a corporate-level plan that articulates a results structure from outputs to impact, it is applied as a framework and widely considered as such by staff. The reason for this lies in the UNICEF modus operandi, which is to support national governments. Programme planning occurs in negotiation with governments around the world. Therefore, UNICEF does not have complete control over its programming at the country-level. The results UNICEF seeks are usually achieved in partnership with governments, as well as with non-governmental organizations and civil society.

The key question is how to determine whether UNICEF is performing well. Therefore if the Strategic Plan transitions to a framework, mechanisms must be found to aggregate results meaningfully at both the country and global levels. Results must remain central to the plan or framework.

Recommendations

The following five recommendations are based on the findings and conclusions presented within this report:

- 1. Reorient the Strategic Plan in line with its application as a framework.** The evaluability assessment finds that the Strategic Plan is articulated logically as a plan but is applied in practice as a framework to better address national development priorities and context. Recognizing this, it is recommended that the next round of strategy development

should more explicitly account for the need for greater flexibility and adaptation (e.g. more flexibility at the level of outputs, reduced use of aggregated results reporting, greater recognition of context-specific variables). This should be fully considered and negotiated with key stakeholders, taking into consideration the UNICEF modus operandi as a highly decentralized organization. The upcoming Mid-Term Review of the Strategic Plan 2014-2017 (MTR) should, in addition to assessing progress in strategy implementation, further examine the implications of using a strategic framework approach for the 2018-2021 period and, accordingly, provide an indication of next steps.

- 2. Newly drafted country programme documents (CPDs) and country programme action plans (CPAPs) should be reviewed to better understand how they relate and contribute to the Strategic Plan.** With substantial regional office support, country offices are making efforts to adjust country programmes for better alignment with the Strategic Plan. Greater alignment of priorities and clarity of results could be facilitated through the review of newly developed CPDs and CPAPs and sharing of lessons. This process should review how well country offices have used determinants analyses and other tools to focus programmes on the rights of the most disadvantaged and excluded children in the country context; and in turn, how well these country-determined foci align with results articulated in the Strategic Plan.
- 3. Clarify results, measurement approaches and indicators for equity.** Equity is a central tenet of the UNICEF mission and is thoroughly integrated into the Strategic Plan. However, there is a lack of clarity in the Strategic Plan on defining results related to equity and indicating how these will be measured. It is not necessary for UNICEF to have just one single definition of equity. Rather, all equity concepts that are employed need to be operationalized with sufficient clarity and rigor to allow measurement and evaluation.
- 4. Conduct a review and adjust, as needed, indicators and systems to measure progress towards the results of the Strategic Plan.** Extensive investments have been made in systems to report on the Strategic Plan indicators. At higher levels of the results chain (i.e. impact and outcome), UNICEF is a central contributor to global data sets, analytical work and methods development. Concerns related to indicator quality at these levels are limited. Performance indicators at the level of outputs are of more variable quality, and despite improvements, there are questions about the systems used to generate data. It is recommended that an assessment of the quality of the output-level performance indicators be conducted with particular attention to those generated through the country office annual reporting system. Any further adjustments to indicators should be based on objective assessment of the performance of this system. The review should be explicit about how results should be measured both at the country and global levels.
- 5. Conduct an exercise to determine the need for and feasibility of moving the organization towards budgeting practices that are more supportive of results-based management that enables better linking of resources and results.** Currently, the Strategic Plan's integrated budget relies on projections of resource availability (i.e. means-based budget). The organization should explore various ways of linking results to resources and capacity.

Resumen ejecutivo

En una etapa inicial de la implementación del Plan Estratégico, la Oficina de Evaluación del UNICEF encargó realizar una valoración de la evaluabilidad del plan. En términos generales, una valoración de la evaluabilidad es un examen de una actividad propuesta para establecer si sus objetivos están definidos adecuadamente y si sus resultados pueden ser objeto de verificación. La valoración de la evaluabilidad proporciona pruebas a la dirección del UNICEF y a las partes interesadas sobre la solidez del plan; el grado en que los indicadores se validan y son medibles; y la medida en que existen sistemas para cuantificar y verificar el desempeño. Este documento presenta los hallazgos y las recomendaciones de la valoración de la evaluabilidad, llevada a cabo desde septiembre de 2014 hasta febrero de 2015. Cabe señalar el enorme esfuerzo que fue necesario para elaborar el Plan Estratégico y los documentos conexos. Se espera que la valoración de la evaluabilidad pueda agregar valor a la hora de mejorar este muy destacable conjunto de materiales.

La valoración de la evaluabilidad se encuadró en torno a cinco objetivos, a saber:

Objetivo 1: Determinar la pertinencia, la claridad y la coherencia del diseño del Plan Estratégico, incluida la integridad de la cadena causal que vincula insumos y actividades con productos, efectos directos e impacto previstos.

Objetivo 2: Determinar la idoneidad y validez de los indicadores, las herramientas y los sistemas para el seguimiento, la medición y la verificación de resultados; y el uso de sistemas y herramientas para valorar la idoneidad y la calidad de los datos que se van a utilizar para el seguimiento, la medición y la evaluación de resultados, incluidas carencias significativas en la cobertura.

Objetivo 3: Determinar los datos y los sistemas para definir el grado en que se pueden evaluar los recursos humanos y económicos frente a los resultados previstos.

Objetivo 4: Ofrecer orientación sobre enfoques de la evaluación del Plan Estratégico con vistas a mejorar su utilidad.

Objetivo 5: Proporcionar pruebas y formular recomendaciones para orientar cualquier mejora necesaria respecto al diseño del Plan Estratégico.

A fin de orientar la valoración de la evaluabilidad se utiliza un modelo analítico centrado en “la evaluabilidad *en principio*”, basado en la lógica interna de un programa según se representa por su teoría del cambio, y “la evaluabilidad *en la práctica*”, según se representa por el enfoque de medición, las herramientas y los indicadores utilizados para reflejar el desempeño y los resultados conseguidos. La valoración de la evaluabilidad tomó como base una combinación de métodos y herramientas para fomentar el rigor y la estandarización al tiempo que integraba información cualitativa para contextualizar y matizar los hallazgos. A

partir de las mejores prácticas identificadas en la documentación disponible, la valoración de la evaluabilidad se sirvió de tablas con criterios de valoración como herramientas para la revisión de documentos de estrategia fundamentales. Asimismo, se obtuvo información adicional crítica mediante entrevistas semiestructuradas con diversos miembros del personal del UNICEF a nivel de sede, oficinas regionales y país. Se visitaron dos oficinas regionales con el propósito de responder preguntas sobre el programa regional y cómo trasladar el Plan Estratégico a programas nacionales. Por último, varias preguntas exigieron el examen de sistemas internos para la presentación de informes sobre resultados y recurrieron a debates sobre elementos de sistemas con el personal y la elaboración de mapas de origen para identificar el flujo direccional de los datos de los resultados desde el origen hasta la sede.

De la valoración de la evaluabilidad surgieron diez hallazgos clave. Estos hallazgos se resumen a continuación junto con implicaciones y recomendaciones.

Hallazgo clave 1: Este Plan Estratégico se percibe ampliamente como el más coherente y sólido desde el punto de vista técnico que el UNICEF ha elaborado hasta la fecha.

Hallazgo clave 2: El Plan Estratégico es pertinente respecto al mandato organizacional y a los acuerdos internacionales con los que se ha comprometido el UNICEF, además de estar en consonancia con los Objetivos de Desarrollo Sostenible.

Hallazgo clave 3: La valoración de la evaluabilidad muestra que el Plan Estratégico del UNICEF sirve de marco general.

Hallazgo clave 4: Una dificultad significativa para la evaluabilidad radica en la ausencia de vínculos lógicos expresados explícitamente en la cadena causal, desde los productos hasta los efectos directos y desde los efectos directos hasta el impacto. Si bien hay declaraciones sobre resultados en cada uno de los niveles de la teoría del cambio, la descripción es muy escasa cuando se trata de explicar cómo un nivel conduce o contribuye a otro. Dicho de otro modo, pese a que existe la *estructura* de una jerarquía de resultados, la ausencia de los enunciados “si-entonces” necesarios para vincular plausiblemente un nivel con otro es flagrante.

Hallazgo clave 5: La valoración concluyó que los indicadores examinados en general funcionaban bien al contrastarlos con criterios tales como la definición del indicador, la existencia de una base de referencia, y la especificidad respecto al resultado previsto. Sin embargo, no puede afirmarse lo mismo de varios indicadores sobre el desempeño incluidos en el Marco de resultados y, en algunos casos, esta falta de solidez puede impedir la evaluación como consecuencia de la mala calidad o una limitada capacidad de interpretación.

Hallazgo clave 6: El prisma de la equidad se aplica a la totalidad del Plan Estratégico; no obstante, el enfoque de la medición de la equidad en el Plan Estratégico es deficiente y se ve obstaculizado adicionalmente por la disponibilidad cuestionable de datos que cuantifiquen el progreso en este ámbito.

Hallazgo clave 7: Los temas transversales dificultan la evaluabilidad.

Hallazgo clave 8: Se han identificado las fuentes y los sistemas de datos para medir indicadores y, o bien se estaban aplicando, o bien se estaban estableciendo en el momento de determinar la evaluabilidad. Se han aprovechado muchas lecciones aprendidas, puesto que estos sistemas de medición de indicadores se implementaron progresivamente y se adoptaron pasos importantes para perfeccionarlos y mejorarlos todavía más. Por cuestiones de tiempo, en la valoración de la evaluabilidad no se pudo examinar un ciclo completo de presentación de informes que incluía estas revisiones. Quedan pendientes algunos ámbitos que son motivo de preocupación; para ellos, las respuestas de SMQ sirven de base para la mayoría de los informes de indicadores sobre el desempeño.

Hallazgo clave 9: Si bien los actuales sistemas de presentación de informes clasifican los compromisos económicos por productos de las oficinas en el país, esferas programáticas, actividades clave, subactividades y estrategias de implementación, los materiales revisados no especificaron de qué manera se realizaría el seguimiento de los flujos de recursos entre niveles, con arreglo a los resultados del Plan Estratégico (es decir, efectos directos y productos). Es importante señalar que no se refleja la necesidad de recursos para lograr los resultados anticipados.

Hallazgo clave 10: Las oficinas en el país, con el firme respaldo de las oficinas regionales, ya han utilizado diversas oportunidades del ciclo de programas nacionales para reflexionar sobre el Plan Estratégico y aplicar ajustes en sus programas. Pese a que no se ha definido en el Plan Estratégico, las oficinas regionales han facilitado la asimilación de elementos del Plan Estratégico en programas nacionales adaptando el Plan Estratégico mundial a las prioridades específicas a escala regional.

La óptica analítica empleada para valorar el Plan Estratégico actual examinó los elementos presentes teniendo en cuenta la relevancia, la plausibilidad y cuestiones estructurales y contextuales. Un plan estratégico es producto de un proceso de planificación que define los objetivos de una organización a medio y largo plazo, así como los recursos necesarios para alcanzarlos. Un plan estratégico sirve para crear e implementar una visión en el seno de una organización y para respaldar a la organización a la hora de tomar decisiones y responder a nuevas cuestiones, necesidades y mandatos. Un plan estratégico debe contener los elementos siguientes: a) objetivos claros, b) dirección y foco de atención, c) cuestiones normativas o principios bien definidos y estándares para medirlos, d) enumeración de las estrategias de implementación, e) armonización de la capacidad y los recursos para que coincidan con los objetivos, el foco de atención y las estrategias de implementación, y f) identificación explícita de resultados e indicadores para medir el progreso. La valoración concluyó que los objetivos y la dirección del Plan Estratégico del UNICEF son claros. Por el contrario, apunta que los estándares para medir las cuestiones normativas no son suficientemente sólidos. La valoración subraya la falta de vínculos expresados específicamente entre productos y efectos directos y entre efectos directos e impacto. Esta

falta de claridad en la relación de cada uno de los efectos directos con el impacto previsto de la organización significa que el Plan Estratégico no se puede utilizar para hacer compensaciones o sopesar decisiones entre esferas de efectos directos en el caso de que surja la cuestión o la necesidad. Asimismo, la relación entre productos y efectos directos es problemática ya que la falta de vínculos identificados específicamente significa que no es posible dictaminar de qué manera los productos van a contribuir a los efectos directos.

La valoración señala que el UNICEF actualmente no puede definir la idoneidad de los recursos para alcanzar los resultados anticipados. Esta dificultad no radica en los sistemas de presentación de informes, sino que deriva del uso de un enfoque para elaborar presupuestos que se basa en los medios (tomando como base los ingresos suministrados) en lugar de en los resultados (ofreciendo un costo agregado del trabajo necesario para lograr los resultados). Esto hace que la organización se encuentre en una situación de cierto riesgo puesto que los resultados que se quieren conseguir no están asociados con necesidades de recursos basadas en pruebas.

La valoración de la evaluabilidad muestra que el Plan Estratégico del UNICEF sirve de marco general. El análisis indica que, si bien el Plan Estratégico se asemeja a un plan a nivel corporativo que establece una estructura de resultados desde los productos hasta el impacto, se aplica como un marco y así es considerado ampliamente por el personal. Esto se explica por el *modus operandi* del UNICEF, que consiste en prestar apoyo a los gobiernos nacionales. La planificación programática se lleva a cabo *negociando* con los gobiernos de todo el mundo. En consecuencia, el UNICEF no tiene el control absoluto de sus programas a nivel de país. Los resultados que quiere lograr el UNICEF normalmente se alcanzan en asociación con gobiernos, además de con ONG y la sociedad civil.

La cuestión fundamental es cómo resolver si el desempeño del UNICEF es correcto. Por lo tanto, si el plan evoluciona hacia un marco, es necesario encontrar mecanismos para agregar los resultados de manera significativa, tanto a nivel nacional como mundial. Los resultados deben seguir siendo el aspecto central del plan o del marco.

Recomendaciones

Las recomendaciones siguientes se basan en los hallazgos y las conclusiones presentados en este informe:

- 1. Reorientar el Plan Estratégico en consonancia con su aplicación como marco.** La valoración de la evaluabilidad ha concluido que el Plan Estratégico se estructura de forma lógica como plan pero que, en la práctica, se aplica como un marco con el cual abordar mejor las prioridades y el contexto de desarrollo nacionales. Reconociendo lo anterior, se recomienda que la siguiente serie de elaboración de estrategias considere más explícitamente la necesidad de contar con una mayor flexibilidad y adaptación (por ejemplo, más flexibilidad a nivel de productos, un uso reducido de informes de resultados agregados, un mayor reconocimiento de las variables específicas del contexto). Todo ello debe estimarse y negociarse plenamente con las partes interesadas clave, valorando el

modus operandi del UNICEF en calidad de organización muy descentralizada. El próximo examen de mitad de período debe, además de determinar el progreso de la implementación de estrategias, examinar adicionalmente las implicaciones de utilizar un enfoque de marco estratégico para el período 2018-2021 y proporcionar una indicación de los pasos siguientes con arreglo a este.

- 2. Se deben revisar los DPP y planes de acción para los programas de los países que se hayan redactado recientemente a fin de comprender mejor su relación con el Plan Estratégico y de qué manera contribuyen a él.** Con el importante apoyo de las oficinas regionales, las oficinas en los países están trabajando a fin de ajustar los programas nacionales para que estén más en consonancia con el Plan Estratégico. Se podría favorecer una mayor armonización de las prioridades y la claridad de los resultados mediante la revisión de nuevos DPP y planes de acción para los programas de los países y el intercambio de lecciones aprendidas. Este proceso de revisión debe tener como propósito comprender si las oficinas en el país han utilizado análisis de factores determinantes y otras herramientas para centrar los programas en los derechos de los niños y las niñas más desfavorecidos y excluidos en el contexto nacional. A su vez, también debe servir para comprender si estos focos de atención determinados por los países coinciden con los resultados enumerados en el Plan Estratégico.
- 3. Esclarecer los resultados, los enfoques de medición y los indicadores de la equidad.** La equidad es un principio fundamental de la misión del UNICEF y como tal está presente en todo el Plan Estratégico. Sin embargo, el Plan Estratégico no es suficientemente claro a la hora de definir los resultados relacionados con la equidad y la manera en la que se van a medir. No es preciso que el UNICEF tenga una única definición de equidad. No obstante, todos los conceptos utilizados en relación con la equidad se deben aplicar con la claridad y el rigor suficientes para permitir la medición y la evaluación.
- 4. Llevar a cabo una revisión y un ajuste, según sea necesario, de los indicadores y sistemas para medir el progreso hacia los resultados del Plan Estratégico.** Se han realizado inversiones importantes en sistemas que permitan informar sobre los indicadores del Plan Estratégico. En niveles superiores de la cadena de resultados (es decir, impacto y efecto directo), el UNICEF contribuye de manera decisiva a los conjuntos de datos mundiales, el trabajo analítico y el desarrollo de métodos. A estos niveles, son pocas las preocupaciones relacionadas con la calidad de los indicadores. Por otro lado, la calidad de los indicadores del desempeño a nivel de productos es más variable y, aunque se observan mejoras, los sistemas utilizados para generar datos son cuestionables. Se recomienda determinar la calidad de los indicadores del desempeño a nivel de productos prestando especial atención a aquellos que se generan mediante el sistema de presentación de informes anuales de las oficinas en el país. Los ajustes adicionales que se puedan aplicar a los indicadores se deben basar en una valoración objetiva del desempeño de dicho sistema. Asimismo, la revisión debe reflejar explícitamente el modo en que se deben medir los resultados tanto a nivel nacional como mundial.
- 5. Llevar a cabo un ejercicio para determinar la necesidad y viabilidad de que la organización avance hacia prácticas para elaborar presupuestos que presten mayor apoyo a una gestión basada en resultados que permita vincular mejor los recursos y los resultados.** Actualmente, el presupuesto integrado del Plan Estratégico se basa en previsiones de disponibilidad de recursos (es decir, presupuesto basado en

medios). La organización debe estudiar diferentes maneras de vincular los resultados a los recursos y la capacidad.

Résumé Analytique

À un stade précoce de la mise en œuvre du Plan stratégique, le Bureau de l'évaluation de l'UNICEF a demandé que l'on procède à une étude d'évaluabilité du plan. De façon générale, une étude d'évaluabilité consiste à revoir une activité proposée afin de déterminer si ses objectifs sont bien définis et si ses résultats sont vérifiables. L'étude d'évaluabilité fournit des éléments factuels à la direction et aux parties prenantes de l'UNICEF concernant : la solidité du plan ; la mesure dans laquelle les indicateurs sont validés et mesurables; et la mesure dans laquelle les systèmes sont en place pour mesurer et vérifier les performances. Ce document présente les conclusions et recommandations de l'étude d'évaluabilité, qui a été menée de septembre 2014 à la fin février 2015. Nous apprécions le fait qu'un énorme effort a été nécessaire pour créer le Plan stratégique et les documents qui s'y rapportent. Nous espérons que cette étude d'évaluabilité permettra d'améliorer encore cet ensemble impressionnant de documents.

L'étude d'évaluabilité était axée sur les cinq objectifs suivants :

Objectif 1 : évaluer l'adéquation, la clarté et la cohérence de la conception du Plan stratégique, y compris l'intégrité de la chaîne de causalité reliant les ressources et les activités avec les extrants, les réalisations et l'impact attendus.

Objectif 2 : évaluer l'adéquation et la validité des indicateurs, outils et systèmes pour le suivi, la mesure et l'évaluation des résultats ; et utiliser des systèmes et des outils pour évaluer la pertinence et la qualité des données qui seront utilisées pour le suivi, la mesure et l'évaluation des résultats, y compris les lacunes significatives dans la couverture.

Objectif 3 : étudier les données et les systèmes afin de déterminer la mesure dans laquelle les ressources humaines et financières peuvent être évaluées par rapport aux résultats attendus.

Objectif 4 : donner des orientations sur les démarches à adopter pour l'évaluation du Plan stratégique en vue d'améliorer son utilité.

Objectif 5 : fournir des preuves et des recommandations en vue d'orienter toute amélioration éventuellement nécessaire concernant la conception du Plan stratégique.

Pour orienter l'étude d'évaluabilité on utilise un modèle analytique centré sur « l'évaluabilité en *principe* », basé sur la logique interne d'un programme tel que représenté par sa théorie du changement, et « l'évaluabilité *dans la pratique* » représentée par l'approche axée sur la mesure, les outils et les indicateurs utilisés pour refléter les performances et résultats obtenus. L'étude d'évaluabilité s'est fondée sur une combinaison de méthodes et outils divers visant à renforcer la rigueur et la standardisation tout en intégrant des informations qualitatives pour placer les conclusions dans leurs contextes et les nuancer. Sur la base des meilleures

pratiques identifiées dans la documentation existante, l'étude d'évaluabilité s'est servie de grilles d'évaluation comme outils pour l'examen des documents de stratégie de base. Des entrevues semi-structurées menées avec tout un échantillon de membres du personnel de l'UNICEF au siège, dans les bureaux régionaux et au niveau des pays, ont permis de recueillir d'autres informations d'importance critique. Deux bureaux régionaux ont été visités afin d'obtenir des réponses aux questions sur le programme régional et la transposition du Plan stratégique dans les programmes de pays. Enfin, il a fallu dans plusieurs secteurs examiner les systèmes internes existants pour rendre compte des conclusions et s'inspirer du débat sur les éléments du système avec le personnel et le développement de cartes d'origine pour identifier le flux directionnel des données de résultats depuis l'origine jusqu'au siège.

L'étude d'évaluabilité a fait apparaître dix principales constatations, résumées ci-dessous avec des implications et recommandations.

Principale conclusion 1 : on admet généralement que ce Plan stratégique est le plus cohérent et le plus solide sur un plan technique jamais mis au point par l'UNICEF.

Principale conclusion 2 : le Plan stratégique est pertinent par rapport au mandat de l'organisation et aux accords internationaux auxquels l'UNICEF est lié et il s'aligne sur les Objectifs de développement durable.

Principale conclusion 3 : l'étude d'évaluabilité constate que le Plan stratégique de l'UNICEF joue un rôle de cadre général.

Principale conclusion 4 : le manque de connexions logiques explicites de la chaîne causale, des extrants aux réalisations et des réalisations à l'impact, représente un problème considérable pour l'évaluabilité. Alors que les déclarations de résultats apparaissent à chaque niveau de la Théorie du changement, la description est très sommaire quand il s'agit d'expliquer comment on passe d'un niveau à un autre ou comment on contribue d'un niveau à un autre. En d'autres termes, alors que la *structure* d'une hiérarchie de résultats est bien présente, les déclarations de « si - alors » nécessaires pour relier plausiblement un niveau à un autre sont largement absentes.

Principale conclusion 5 : l'évaluation a conclu que les indicateurs examinés se sont généralement bien comportés comparé aux critères, notamment, la définition des indicateurs, l'existence d'une base de référence et la spécificité vis-à-vis du résultat attendu. Cependant, quelques indicateurs de performance que l'on trouve dans le cadre de résultats sont plus faibles et, dans certains cas, peuvent même faire obstacle à l'évaluation à cause de leur mauvaise qualité ou de la possibilité limitée de leur interprétation.

Principale conclusion 6 : le prisme de l'équité s'applique à la totalité du Plan stratégique mais l'approche de la mesure d'équité dans le Plan stratégique est problématique et elle est encore affaiblie par la disponibilité douteuse de données pour mesurer les progrès dans ce domaine.

Principale conclusion 7 : les thèmes transversaux posent des problèmes à l'évaluabilité.

Principale conclusion 8 : les sources de données et les systèmes de mesure des indicateurs sont identifiés et ils sont mis en place au moment de l'étude d'évaluabilité ou bien en train d'être mis en place. De nombreux enseignements tirés ont été pris en compte à mesure qu'étaient mis en œuvre ces systèmes de mesure des indicateurs et l'on a pris des mesures importantes pour les affiner et les améliorer. Compte tenu du calendrier, l'étude d'évaluabilité n'a pas été en mesure d'examiner la totalité du cycle de présentation des résultats qui comprenait ces révisions. Certains sujets de préoccupation demeurent à propos des réponses SMQ qui servent de base à la majorité des communications sur les indicateurs de performance.

Principale conclusion 9 : bien que les systèmes actuels de rapports classent les engagements financiers en fonction des extraits des Bureaux de pays, de leurs domaines de programmes, de leurs principales activités, sous-activités et stratégies de mise en œuvre, les documents examinés ne précisaient pas la façon dont les flux de ressources, à différents niveaux, seraient suivis en fonction des résultats du Plan stratégique (c.-à-d. les réalisations et les extraits). Il importe de noter que l'on ne précise pas les ressources nécessaires pour atteindre les résultats attendus.

Principale conclusion 10 : les Bureaux de pays, avec l'appui vigoureux des Bureaux régionaux, ont déjà utilisé un certain nombre de possibilités dans le cycle de programme de pays pour étudier le Plan stratégique et procéder à des ajustements dans leurs programmes. Bien que cela ne soit pas précisé dans le Plan stratégique, les Bureaux régionaux ont facilité l'adoption d'éléments du Plan stratégique dans les programmes de pays en adaptant le Plan stratégique mondial à diverses priorités spécifiques à l'échelle régionale.

Le prisme analytique utilisé pour évaluer le Plan stratégique actuel a permis d'examiner les éléments en place, en prenant en compte la pertinence, la plausibilité, les problèmes structurels et contextuels. Un plan stratégique est le produit d'un processus de planification qui définit les objectifs à moyen et long terme d'une organisation, ainsi que les ressources nécessaires pour les atteindre. Un plan stratégique permet d'établir et de mettre en œuvre une vision au sein d'une organisation et d'aider l'organisation à faire des choix et à répondre aux questions, besoins et mandats en train d'apparaître. Un plan stratégique doit contenir les éléments suivants : a) des objectifs clairs ; b) une orientation et des priorités ; c) des principes bien définis et des questions normatives ainsi que des normes pour les mesurer ; d) l'articulation des stratégies de mise en œuvre ; e) l'alignement des capacités et des ressources sur les objectifs, priorités et les stratégies de mise en œuvre ; et f) l'identification explicite des résultats et des indicateurs pour mesurer les progrès réalisés. L'évaluation conclut que les objectifs du Plan stratégique de l'UNICEF et son orientation sont clairs. Elle note que les normes utilisées pour mesurer les questions normatives sont faibles. L'évaluation met en évidence le manque de liens spécifiquement déclarés entre les extraits et les réalisations, ainsi qu'entre les réalisations et l'impact. Ce manque de clarté sur la façon dont

chaque réalisation est reliée à l'impact souhaité de l'organisation signifie que le Plan stratégique ne peut être utilisé pour faire des compromis ou peser des décisions dans les secteurs des réalisations au cas où la question se poserait ou en cas de besoin. La relation entre les extrants et les réalisations est tout aussi problématique puisque le manque de liens spécifiquement identifiés signifie qu'il n'est pas possible d'établir comment les extrants contribuent aux réalisations.

L'évaluation indique que l'UNICEF ne peut actuellement déterminer l'adéquation des moyens à mettre en œuvre pour atteindre les résultats escomptés. Ce problème ne tient pas aux systèmes d'établissement des rapports, mais résulte de l'utilisation d'une approche budgétaire qui est axée sur les moyens (basée sur les revenus fournis) au lieu d'être axée sur les résultats (fournir un coût total des travaux nécessaire pour atteindre les résultats). Cela place l'organisation en situation de risque car les résultats à obtenir ne sont pas associés à des demandes en matière de ressources qui soient fondées.

L'étude d'évaluabilité montre que le Plan stratégique de l'UNICEF constitue un cadre général. L'analyse note que le Plan stratégique joue le rôle d'un cadre de travail et que c'est ainsi qu'il est considéré généralement par le personnel, même s'il ressemble à un plan d'entreprise qui présente une structure de résultats des réalisations jusqu'à l'impact. Cela s'explique par le mode de fonctionnement de l'UNICEF qui consiste à épauler les gouvernements nationaux. La planification des programmes a lieu *lors des négociations* avec les gouvernements du monde entier. Par conséquent, l'UNICEF ne maîtrise pas totalement sa programmation au niveau des pays. Les résultats que l'UNICEF cherche sont habituellement réalisés en partenariat avec les gouvernements ainsi qu'avec les ONG et la société civile.

La question clé est de savoir comment déterminer si l'UNICEF fonctionne bien. Donc, si le plan se transforme en un cadre, il est impératif de trouver des mécanismes permettant de totaliser les résultats de manière utile tant au niveau du pays qu'au niveau mondial. Les résultats doivent être au centre du plan ou du cadre de travail.

Recommandations

Les recommandations suivantes se fondent sur les constatations et les conclusions présentées dans ce rapport :

- 1. Réorienter le Plan stratégique pour qu'il coïncide avec son application en tant que cadre de travail.** L'étude d'évaluabilité a constaté que le Plan stratégique se structure logiquement comme un plan, mais qu'il est appliqué dans la pratique comme un cadre pour mieux aborder les priorités et contextes nationaux de développement. Compte tenu de cette observation, on recommande que le prochain cycle de développement de la stratégie tienne plus explicitement compte de la nécessité d'une plus grande flexibilité et d'une meilleure adaptation (par exemple plus de souplesse au niveau des extrants, une utilisation moins fréquente des agrégats de résultats, une reconnaissance accrue des variables spécifiques à certains contextes). Cela devrait être considéré dans sa totalité et négocié avec les principales parties prenantes, en tenant compte du fonctionnement très

décentralisé de l'UNICEF. Outre l'évaluation des progrès obtenus dans la mise en œuvre de la stratégie, le prochain examen à mi-parcours doit examiner les conséquences de l'utilisation d'une approche avec un cadre stratégique pour la période 2018-2021 et fournir donc une indication des prochaines étapes.

- 2. Les nouveaux descriptifs de programmes de pays et les plans d'action pour la mise en œuvre des programmes de pays devraient être revus afin de mieux comprendre leur rapport avec le Plan stratégique et comment ils y contribuent.** Avec l'appui important du Bureau régional, les bureaux de pays s'efforcent d'ajuster les programmes de pays pour mieux les aligner sur le Plan stratégique. On pourrait faciliter une meilleure harmonisation des priorités et rendre les résultats plus clairs en examinant les nouveaux descriptifs de programmes de pays et plans d'action pour la mise en œuvre des programmes de pays et en partageant les enseignements tirés. Ce processus d'examen devrait avoir pour but de voir si les bureaux de pays ont utilisé les analyses de facteurs déterminants et d'autres outils pour axer leurs programmes sur les droits des enfants les plus défavorisés et marginalisés dans le contexte du pays ; et également de voir si ces priorités déterminées par les pays coïncident avec les résultats énoncés dans le Plan stratégique.
- 3. Clarifier les résultats, les méthodes de mesure et les indicateurs concernant l'équité.** Principe central de la mission de l'UNICEF, l'équité est complètement intégrée au Plan stratégique. Cependant, le Plan stratégique souffre d'un manque de clarté concernant la définition des résultats relatifs à l'équité et la façon dont ils seront mesurés. Il n'est pas nécessaire que l'UNICEF n'ait qu'une seule définition de l'équité. Cependant, tous les concepts d'équité qui sont employés doivent être opérationnalisés avec suffisamment de clarté et de rigueur pour permettre les mesures et l'évaluation.
- 4. Procéder à un examen et ajuster, le cas échéant, les indicateurs et systèmes pour mesurer les progrès vers les résultats du Plan stratégique.** D'importants investissements ont été réalisés dans les systèmes de rapport sur les indicateurs du Plan stratégique. Aux niveaux supérieurs de la chaîne des résultats (c'est-à-dire impact et réalisations), l'UNICEF contribue de manière décisive à la collecte d'ensembles de données mondiales, aux travaux d'analyse et à l'élaboration de méthodes. On ne s'inquiète guère de la qualité des indicateurs à ces niveaux. Les indicateurs de performance au niveau des extrants sont de qualité plus variable et, malgré les améliorations, on peut s'interroger sur les systèmes utilisés pour générer des données. Il est recommandé de procéder à l'évaluation de la qualité des indicateurs de performance au niveau des extrants en s'intéressant en particulier à ceux qui sont générés par le système d'établissement de rapports annuels par les bureaux de pays. Tout nouvel ajustement des indicateurs devrait reposer sur une évaluation objective de la performance du système d'établissement de rapports annuels par les bureaux de pays. L'examen doit être explicite sur la façon dont les résultats devraient être mesurés tant aux niveaux national que mondial.
- 5. Organiser un exercice visant à déterminer la nécessité et la possibilité d'orienter l'organisation vers des pratiques de budgétisation qui soutiennent davantage une gestion axée sur les résultats et qui permettent de mieux lier les ressources aux résultats.** Actuellement, le budget intégré du Plan stratégique s'appuie sur des projections concernant la disponibilité des ressources (c.-à-d. un budget axé sur les moyens). L'organisation devrait explorer diverses façons de lier les résultats aux ressources et aux capacités.

1. Introduction and background

The UNICEF Evaluation Office conducted an evaluability assessment of the UNICEF Strategic Plan, 2014-2017 during an early stage of its implementation. Broadly stated, an evaluability assessment is a review of a proposed activity to determine whether its objectives are adequately defined and its results are verifiable. The purpose of this document is to present the findings and recommendations of the evaluability assessment, which was conducted from September 2014 to February 2015. In the context of the UNICEF Strategic Plan, this evaluability assessment provides evidence to UNICEF management and stakeholders regarding how robust the plan is, the degree to which indicators are validated and measurable and the extent to which systems are in place to measure and verify performance. As such, this assessment is distinctly different from an evaluation and should not be considered as an evaluation. Conducting the exercise in the first year of the strategy cycle permits UNICEF to strengthen, where necessary, the systems needed to reliably measure, monitor and evaluate progress toward results. The evaluability assessment will also identify issues for the MTR and the formulation of the next strategy. Accordingly, the main audience for the evaluability assessment is senior management within UNICEF.

1.1 The UNICEF Strategic Plan, Theory of Change and results framework

As mandated by the United Nations General Assembly, UNICEF advocates for the protection of children's rights, helps meet their basic needs and expands their opportunities to reach their full potential. To guide the organization in implementing this mission, UNICEF developed a Strategic Plan, 2014-2017. A strategic plan is a product of a planning process that defines an organization's medium- to long-term goals, as well as the resources to achieve them. A strategic plan serves to establish and implement a vision within an organization and to support the organization in making choices and responding to emerging issues, needs and mandates.

The Strategic Plan, 2014-2017 replaced the UNICEF medium-term strategic plan 2006-2013 (MTSP), which was developed in 2005 and extended twice in duration by the Executive Board. Indeed, the Strategic Plan, 2014-2017 was developed based on the *Report on the end-of-cycle review of the medium-term strategic plan 2006-2013* (Strategic Plan, page 3) and a commitment to results-based management.¹ The Strategic Plan's development process was highly consultative. Elements of the process included directives to synchronize strategy cycles and, to some extent, to synchronize structures with the United Nations Development Programme (UNDP), the United Nations Population Fund (UNFPA) and the United Nations Entity for Gender Equality and the Empowerment of Women (UN Women), as well as to align with the Quadrennial Comprehensive Policy Review. Multiple rounds of consultations were carried out with the Executive Board, as well as with Member States, other United Nations agencies and international non-governmental organizations. In addition, the Strategic Plan was guided by a peer review group of experts on Results-Based Management drawn from

¹ This effort was positively noted in a United Nations system-wide review of strategic planning as an example of a means to increase "efficiency and effectiveness in the use of resources and in facilitating the work of Member States by consolidating planning and reporting on issues of common concern across the system." United Nations, *Strategic Planning in the United Nations System*, United Nations Joint Inspection Unit, Geneva, 2012.

United Nations agencies (the Joint United Nations Programme on HIV/AIDS (UNAIDS), UNDP, UNFPA, the United Nations Educational, Scientific and Cultural Organization (UNESCO), UN Women, the World Food Programme (WFP) and the World Health Organization (WHO)) and Member States (Australia, Canada, China, Djibouti, Kenya, the Netherlands, Norway, Pakistan and the United Kingdom).² The Executive Board approved the Strategic Plan in September 2013.

Some of the major ways in which the Strategic Plan differs from the MTSP include: its structure, the departure from thematic areas to more clearly sectoral outcome areas, the integration of humanitarian action and gender across programme elements, the creation of an outcome area on social inclusion, and a sharper focus on resilience. Through both its timing and development process, the Strategic Plan 2014-2017 reflects UNICEF positioning in a changed environment, including the emergence of macro-level trends that are having significant impacts on children.

The Strategic Plan was fundamentally informed by the UNICEF equity refocus. It opens with the following overarching statement of purpose, framing the entire endeavour from an equity perspective: “The fundamental mission of UNICEF is to promote the rights of every child ... The equity strategy, emphasizing the most disadvantaged and excluded children and families, translates this commitment to children’s rights into action. For UNICEF, equity means that all children have an opportunity to survive, develop and reach their full potential, without discrimination, bias or favouritism. To the degree that any child has an unequal chance in life – in its social, political, economic, civic and cultural dimensions – her or his rights are violated ... This is why the focus on equity is so vital.”³

The Strategic Plan is designed to enable the organization to contribute to the impact of “the realization of the rights of all children, especially the most disadvantaged.”⁴ This impact will be achieved in part through the contributions of seven programmatic outcomes,⁵ which reflect institutional and behavioural changes, as per United Nations Development Group (UNDG) guidelines:

1. Health: Improved and equitable use of high-impact maternal, newborn and child health interventions from pregnancy to adolescence and promotion of healthy behaviours;
2. HIV and AIDS: Improved and equitable use of proven HIV prevention and treatment interventions by children, pregnant women and adolescents;
3. Water, sanitation, and hygiene (WASH): Improved and equitable use of safe drinking water, sanitation and healthy environments, and improved hygiene practices;

² The Peer Review Group list is available at: UNICEF, ‘Strategic Plan 2014-2017’, <<http://www.unicef.org/strategicplan/>>, accessed 12 August 2015.

³ United Nations Children’s Fund, ‘The UNICEF Strategic Plan, 2014-2017’, UNICEF Executive Board second regular session 2013, Item 4 of the provisional agenda, New York, 11 July 2013, <http://www.unicef.org/strategicplan/files/2013-21-UNICEF_Strategic_Plan-ODS-English.pdf>, accessed 12 August 2015, p. 1.

⁴ Ibid., p. 5.

⁵ Ibid., pp. 6-8.

4. Nutrition: Improved and equitable use of nutritional support and improved nutrition and care practices;
5. Education: Improved learning outcomes and equitable and inclusive education;
6. Child protection: Improved and equitable prevention of and response to violence, abuse, exploitation and neglect of children;
7. Social inclusion: Improved policy environment and systems for disadvantaged and excluded children, guided by improved knowledge and data.

In turn, each one of the outcomes will be achieved in part through the combined results of six common outputs:⁶

1. Demand: Enhanced support for children, families and communities to promote knowledge, behaviour change, demand for services and opportunities for participation;
2. Access: Increased national capacity to ensure availability of and access to services and to strengthen systems;
3. Enabling environment: Strengthened political commitment and national capacity to legislate, plan and budget for children;
4. Humanitarian: Increased country capacity and delivery of services to protect and provide for children in humanitarian settings;
5. Gender equality and human rights: Increased capacity of governments and partners, as duty bearers, to identify and respond to specific challenges related to the protection and promotion of the rights of children and gender equality;
6. Global and regional programmes: Enhanced enabling environment at global and regional levels to realize child rights.

The Strategic Plan document is supplemented by the detailed Revised Supplementary Programme Note on the Theory of Change for the UNICEF Strategic Plan, 2014-2017, the final results framework of the UNICEF Strategic Plan, 2014-2017 and the UNICEF integrated budget, 2014-2017. The results framework responds to a recommendation of the Quadrennial Comprehensive Policy Review for a detailed document that demonstrates the complete results chain at output, outcome and impact levels and includes indicators with baselines, milestones and targets for monitoring. Similarly, the Theory of Change document unpacks the results chain, seeking to explain *why* UNICEF selected these strategic priorities and *how* UNICEF envisions the causal pathways will achieve the results set out in the results framework. In other words, it provides a rationale for the changes that UNICEF seeks to accomplish and contribute to, and describes how the organization believes that change will occur.

⁶ The output is a capacity change for the specified rights holder or duty bearer: a new skill, ability or the availability of a new product or service.

2. Approach, objectives and methods

This chapter briefly summarizes the approach, objectives, questions, methods of data collection and mode of analysis.

2.1 Approach of the evaluability assessment

The overall approach was informed by internationally accepted standards and best practices for conducting evaluability assessments. The Strategic Plan was neither benchmarked against the strategic plans of other agencies nor compared to an abstract standard. The evaluability assessment should provide a baseline against which any later improvements or adjustments can be assessed.

A team of two independent consultants managed by the UNICEF Evaluation Office carried out the evaluability assessment. An internal reference group provided direction and feedback on materials as they were developed. The evaluability assessment required close collaboration with UNICEF staff at Headquarters and regional and country offices. The assessment was conducted in three phases: 1) an inception phase; 2) a development phase; and 3) validation phase. Further details on the approach of the evaluability assessment can be found in the inception report in Annex 1, which includes the terms of reference.

2.2 Objectives

The evaluability assessment was framed around five objectives (*see Table 1*). While the first three of these objectives and the related questions aimed to generate the required evidence, Objectives 4 and 5 focused on the interpretation and synthesis of that evidence in order to make recommendations.

2.3 Methods

The evaluability assessment relied on a mix of methods and tools to boost rigour and standardization. Qualitative information was particularly important to contextualize and nuance the findings. The range of methods used throughout the development phase is laid out in an evaluation matrix, which appears in the inception report in Annex 1. For each question, the evaluation matrix specifies data sources, data collection methods and analytical approaches. The matrix provides the agreed basis from which all conceptual aspects of the assignment were drawn.

An analytical construct that was used to guide the evaluability assessment centred on evaluability in principle and in practice. Evaluability in principle examines the internal logic of a programme, as represented by its theory of change, and evaluability in practice examines the measurement approach, tools and indicators used to reflect performance and results achieved. Based on best practices identified in the literature, the evaluability assessment utilized rubrics to examine both the in-principle and in-practice aspects of evaluability. The section below summarizes information that appears in more detail in the inception report in Annex 1.

Table 1: Evaluability assessment objectives and key questions

<p>Objective 1: Assess the relevance, clarity, and coherence of the Strategic Plan's design, including the integrity of the causal chain linking inputs and activities with expected outputs, outcomes and impact.</p>
<ul style="list-style-type: none"> • How relevant is the Strategic Plan in relation to UNICEF charter documents? • To what extent do the Strategic Plan, Theory of Change and results framework present a clear and comprehensive causal chain? • How plausible and testable are the Theory of Change and associated results, indicators and activities? • What are the issues pertaining to the evaluability, in principle, of crosscutting priorities and normative principles? • How well have key assumptions, risks and mitigation strategies been specified?
<p>Objective 2: Assess the adequacy and validity of the indicators, tools and systems for monitoring, measuring and verifying results; and of systems and tools to assess the adequacy and quality of data to be used in monitoring, measuring and evaluating results, including any significant gaps in coverage.</p>
<ul style="list-style-type: none"> • Are indicators in place? Have the indicators been defined (e.g. numerators and denominators) with clearly understood standards? • Are baselines in place for the indicators? • Have specific sources of information been identified for the indicators and are systems in place to verify those sources? • Is there a data generating system in place with defined responsibilities, sources and periodicity? • Are measures, tools and mechanisms in place to measure crosscutting priorities and normative principles? • What are the opportunities and challenges related to translating the Strategic Plan into country programmes of cooperation?
<p>Objective 3: Assess data and systems to determine the degree to which human and financial resources can be evaluated against the expected results.</p>
<ul style="list-style-type: none"> • Are data and systems in place to allow UNICEF to link resource allocations to intended results? • Are the data and systems in place to allow UNICEF to assess the adequacy of resources to achieve intended results?
<p>Objective 4: Provide guidance on approaches to the evaluations of the Strategic Plan with a view to enhancing its utility.</p>
<ul style="list-style-type: none"> • Is there an evaluation, research, learning agenda, strategy or plan in place for each outcome? • Is there an evaluation, research, learning agenda, strategy or plan in place for each crosscutting theme and normative principle? If so, do these fit into any overarching learning strategy for the Strategic Plan? • How might the utility of planned evaluations be enhanced? If so, do these fit into any overarching learning strategy for the Strategic Plan?
<p>Objective 5: Provide evidence and recommendations to guide any enhancements that may be required regarding the design of the Strategic Plan.</p>
<ul style="list-style-type: none"> • How can the design of the Strategic Plan be substantially enhanced based on the evidence assembled during the evaluability assessment?

Evaluability in principle was examined using a template with criteria and questions outlined below in Table 2, which is derived from the evaluability matrix. These templates were completed for each outcome area through review of the relevant sections of the Strategic Plan, Theory of Change and results framework against each of the evaluability criteria. A short narrative was prepared based on the stated logic and structured assessment about how well

each criterion was met with reference to specific content in the documents. The systematic narratives generated from the use of these tools combined with interview data were discussed with outcome owners and synthesized into the evaluability assessment report narrative.

Table 2. Criteria to assess evaluability in principle

Evaluability criteria	Key questions
Relevance and clarity	<p>Are the Theory of Change components clearly relevant to the needs of the target group, as identified by any form of situation analysis, baseline study or other evidence and argument?</p> <p>Is the intended beneficiary group clearly identified?</p> <p>Are the elements of the Theory of Change clearly identified and are the proposed steps towards achieving these clearly defined?</p>
Plausibility and testability	<p>Is there a continuous causal chain connecting outputs and outcomes to the final impact of concern?</p> <p>Is it possible to identify which linkages in the causal chain will be most critical to the success of the strategy, and thus the focus of evaluation questions?</p>
Context	<p>Have assumptions about the roles of other actors outside of UNICEF been made explicit (both enablers and constrainers)?</p> <p>Are there plausible plans to monitor these in any practicable way?</p>

Evaluability in practice was centred on the quality of performance indicators and the systems to generate those indicators. Meeting this objective required careful review of the results framework, supplementary and revised indicators, and indicator sheets that were developed for each of the outcome areas. The criteria used to systematically examine the dimensions of indicator quality appear in Table 3. As with Objective 1 methods, these criteria have been extracted from the literature on evaluability assessments and theories of change.

Two types of criteria were used in this assessment. The first were those criteria amenable to being objectively gauged: system or source identified; baseline specified; target specified; quality; pertains to crosscutting priority. The second type was categorized based on the consultants' knowledge of proper indicator construction (definition and reliability) and the logic of the relationship of each indicator to its respective result statement (specificity). The final criterion, sensitivity to change, considered the suitability of the indicator for enabling actions (managing) based on results.

The approach was differentiated based on indicator result level. At the level of impact and outcome indicators, all impact and outcome measures were examined using a light review. Light review was defined as a relatively quick assurance that these indicators largely agreed with inter-agency working groups or other normative bodies with well-established data sources and quality. This approach was justified by the nature of these indicators. It also reflects that results are achieved through the efforts of many actors and are well beyond

Table 3: Evaluability in practice: Dimensions of indicator quality

Criterion	Answers the question(s)	Rubric
Baseline specified	Has a baseline measure been provided for the indicator, including any necessary stratification?	Yes/no
Target specified	Has a target value for the indicator been provided, including for any necessary stratification?	Red = no description of how target was set or no target Yellow = target set based on expert opinion with no further description Green = target set through transparent process and/or aligned with global or regional commitment
System or source identified	Has a specific source or system been identified to collect the indicator?	Yes/no
Pertains to crosscutting	Does the indicator pertain to equity, humanitarian situations or gender equality?	Equity (E) Humanitarian (H) Gender equality (G)
Definition	Has the indicator been defined with, for example, a numerator and denominator where appropriate? Are there standards defined that gauge acceptable quantities and qualities?	(i) Not defined, or defined with one or more elements of a highly subjective manner; (ii) All indicator components clearly defined with some elements of subjectivity remaining; (iii) All indicator components clearly and objectively defined; little or no room for subjectivity.
Reliability of measurement	How reliably can the indicator be measured given specifications in the results framework and indicator sheets? Do different reviewers gauge progress in the same manner? Will the reported value be the same regardless of who collects the data?	(i) Poorly constructed, framed in immeasurable manner or measurability unclear even from specifications provided; (ii) Indicator with weak specification or source or dependent on subjective assessment (e.g. by country office staff member); (iii) Clear or standard indicator easily measured using well-established practice.
Specificity	Does the indicator directly reflect the type of result to be achieved at the impact, outcome or output level? Will the indicator capture what is expected to happen?	(i) Mismatch between type of results expected and indicator (impact, outcome, output); (ii) Possible relationship to type of results anticipated (impact, outcome, output); (iii) Indicator matches type of result anticipated (impact, outcome, output).

UNICEF's direct accountability. A more substantive review was conducted for the output level indicators because UNICEF is more directly accountable for the output level results. However, because the overall number of output indicators is substantial (105 indicators not including those considered supplemental), 26 output indicators were selected to keep the overall evaluability assessment tractable. Indicators were selected from each outcome area and from each output area. The selection process is described in the inception report in Annex 1.

In addition to the criteria described in Table 3, consideration was also given to whether established mechanisms are in place to check and assure data quality for the indicators. Annex 2 contains the document review/extraction templates used in the exercise and discussion guides.

Where needed, additional methods supplemented and reinforced findings based on the use of rubrics. In general, this was done through a combination of document review, extensive stakeholder interviews and systems mapping. These methods are described briefly below.

Certain questions to be answered by the evaluability assessment called for document review and content analysis without the use of rubrics and templates. Where possible, document review and content analysis were used to supplement more systematic methods. This material is footnoted and incorporated throughout this report. A list of all documents reviewed appears in Annex 3.

Semi-structured interviews with a range of UNICEF staff members from Headquarters and regional and country offices provided an additional critical information base. At Headquarters, interviews were conducted during both inception and development phases with individuals and groups across divisions (i.e. from the Division of Data, Research and Policy, Programme Division, Private Fundraising and Partnerships Division and the Division of Communication), as well as within each outcome or crosscutting area. At the regional office level, interviews were conducted with staff members, including directors, deputy directors and chiefs of planning and monitoring and evaluation. In each case, regional office staff members were queried for input on candidate country offices for inclusion, considering the country programme cycle and, most importantly, recent events such as the preparation of a CPD, MTR or strategic moment of reflection. At the country office level, representatives, deputy representatives, and monitoring and evaluation and planning staff were prioritized for interviews. In total, 74 semi-structured interviews were conducted in person and via phone/Skype with stakeholders from Headquarters (N=34), regional offices (N=28) and country offices (N=12).

For each respondent category, interview guides were developed and, when possible, shared in advance of the interview. All interview responses, whether recorded or transcribed, were uploaded into qualitative data analysis software (QDA Miner Lite software) and coded by theme. Those responses, by theme, were then used to augment data collected through other methods. A list of individuals interviewed appears in Annex 4.

Two regional offices were visited in order to answer questions on the regional programme and the translation of the Strategic Plan into country programmes. Based on selection criteria described in the inception report, the Central and Eastern Europe and the Commonwealth of Independent States (CEE/CIS) and the East and Southern Africa regions were selected. These visits allowed first-hand observation of the procedures and processes involved in the translation of the Strategic Plan into country programmes and the opportunity to engage senior managers and regional advisors more directly over the course of several days.

The terms of reference included an objective that would have required an assessment of the extent to which human and financial resource allocations are adequate to achieve expected results. Based on interviews carried out with UNICEF staff during the inception phase and discussion with the reference group, the objective was re-framed to focus on the *systems* needed to assess resources in relation to the results to be achieved. These questions required the examination of internal systems used for results reporting. In these cases, it was necessary to discuss systems elements at length with staff members responsible for conceptualizing, updating and maintaining these systems within UNICEF and then to develop source maps to identify the expected directional flow of results data from source to Headquarters.⁷ Mapping software (i.e. XMind) was used to create the source maps.

This report presents findings on evaluability in principle (Chapter 3) followed by evaluability in practice (Chapter 4). Chapter 5 presents the findings on the translation of the Strategic Plan to the regional and country levels. Chapter 6 draws from the evaluability assessment's findings to offer recommendations for future evaluations and Chapter 7 includes implications and recommendations for UNICEF.

2.4 Limitations of the evaluability assessment

The evaluability assessment was a focused exercise. Time and budgetary constraints forced trade-offs between depth and breadth of coverage. For example, it was not feasible to conduct a detailed examination of every output indicator contained in the updated results framework. In addition, the evaluability assessment did not set out to examine or make comparisons with the strategic plans of other agencies. Interviews were not sought with members of the Executive Board.

The evaluability assessment's utility depended in part on its timely implementation, to allow adjustments to be made in the Strategic Plan. The evaluability assessment thus focused on the Headquarters level, with coverage of all regional offices via two site visits and phone/Skype interviews and coverage of a sample of country offices via phone/Skype interviews. As a result, the evaluability assessment offers a snapshot of experiences translating the Strategic Plan to the country level to date, but could not cover the breadth of that experience (e.g. across differing country contexts) or depth (e.g. through attending country office planning meetings or observing a country office complete the SMQs).

Likewise, timing and timeliness concerns forced the examination of some systems and structures before they have become fully operational. During the evaluability assessment, work was ongoing on central elements of performance monitoring, including the detailed indicator sheets, the Results Assessment Module (RAM) and the SMQs. As the annual reporting process was launched in November 2014, it was not possible to report on how well these systems performed in this report. In addition, organizational re-structuring undertaken in 2014 meant that some important new units (e.g. the Field Results Group) were not yet fully operational.

⁷ These include: results framework; RAM; country operations portal; detailed indicator sheets; and the Theory of Change.

3. Evaluability of the UNICEF Strategic Plan in principle

The first dimension examined was evaluability in principle, which refers to the internal logic of the Strategic Plan. This chapter addresses the questions associated with the evaluability assessment's first objective: to assess the clarity, relevance and coherence of the Strategic Plan's design, including the integrity of the causal chain linking inputs and activities with expected outcomes and goals. As described in the Chapter 2 section on methods, the approach utilized examined several key criteria related to evaluability in principle, as found primarily in the Theory of Change document.

The criteria covered in this chapter include relevance, plausibility and testability and context (assumptions). Findings are further illustrated with qualitative information drawn from interviews.

As outlined above, the Strategic Plan was developed in a consultative manner, with many stakeholders were involved in its development. Marking a clear departure from the MTSP, there has been an attempt to chart causal pathways and assumptions through the development of a Theory of Change document. With near uniformity, staff members felt that although there was room for improvement, the logical chain involved in the Strategic Plan was far superior to any utilized in the past. As will be further described below, the organization mounted a rapid response to the realization that more resources and measurement support were needed to ensure better quality data from country office reporting on Strategic Plan performance indicators. Moreover, with its investments in a theory of change, the organization is well positioned to advance its results agenda in a coherent manner.

This is the first time there has been an attempt to chart causal pathways and assumptions through the development of a Theory of Change document. With near uniformity, staff felt that the logical chain involved in the Strategic Plan was far superior to any utilized in the past.

Key finding 1: This Strategic Plan is widely viewed as the most coherent and technically sound that UNICEF has developed to date.

3.1 Relevance

The relevance concept deals with whether programme objectives remain appropriate over time, particularly as conditions or circumstances change.⁸ This section provides a brief consideration of evaluability in principle, based on the criterion of relevance.

For the UNICEF Strategic Plan, relevance is linked to an underlying set of analytical processes and key documents that set the stage for both strategy development and eventually evaluation. These are: a) the organizational mandate and international agreements to which UNICEF is bound; b) the end-of-cycle review of the previous MTSP 2006-2013; and c) the

⁸ Organisation for Economic Co-operation and Development, *Glossary of Key Terms of Evaluation and Results Based Management*, OECD, Paris, 2002.

global-level situation assessments and analyses created to inform strategy development. These three sources are reviewed briefly below.

Relevance of the Strategic Plan at the impact level and for each of the seven outcomes can be assessed against the UNICEF mandate, other foundational documents and international agreements and accords. Documents for this purpose include the Convention on the Rights of the Child, the Core Commitments for Children in Humanitarian Action, the Millennium Development Goals and the SDGs.⁹ Assessment of the Strategic Plan against these documents provides insight into its relevance as a document that bridges the period immediately before and following the 2015 Millennium Development Goals events.

Table 4: Relevance of the Strategic Plan to existing commitments and the Sustainable Development Goals

Convention on the Rights of the Child (1990)	Millennium Development Goals (2000)	Core Commitments for Children in Humanitarian Action (2010)	UNICEF Strategic Plan (2014-2017)	Sustainable Development Goals (2015-2030)
Article 24	Goals 4 and 5	Health strategic result	Health outcome	Goal 3: Ensure healthy lives and promote well-being for all at all ages
Article 24	Goal 6	HIV/AIDS strategic result	HIV/AIDS outcome	Goal 3: Ensure healthy lives and promote well-being for all at all ages
Article 24	Goal 7	WASH strategies result	WASH outcome	Goal 6: Ensure availability and sustainable management of water and sanitation for all
Article 24	Goal 1	Nutrition strategic result	Nutrition outcome	Goal 2: End hunger, achieve food security and improved nutrition and promote sustainable agriculture
Article 28	Goal 2	Education sector strategic result	Education outcome	Goal 4: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all
Articles 34-39, among others	--	Child protection sector strategic result	Child protection outcome	Goal 16: Promote peaceful and inclusive societies for sustainable development (...)
--	Goal 1	--	Social inclusion outcome	Goal 1: End poverty in all its forms everywhere

As seen in Table 4, the relevance of the Strategic Plan can be assessed at a macro level against the agreements mentioned above and the SDGs. These commitments and agreements have extensive goals and objectives, many of which are crosscutting in nature.

⁹ United Nations, *Report of the Open Working Group of the General Assembly on Sustainable Development Goals*, A/68/970, United Nations, New York, 12 August 2014.

Table 4, with its one-to-one correspondence, is certainly a simplification.

Though the SDGs were not developed at the time that the Strategic Plan was developed, in terms of forward-looking goals, there is a high degree of correspondence between the SDGs and the Strategic Plan's core components. This is the case both in terms of the goals laid out in Figure 1 and the goals related to achieving gender equality and reducing inequalities within and between countries. Further congruence is seen at the level of SDG sub-goals in priority areas such as building resilience and substantially increasing the availability of high-quality, timely and reliable and disaggregated data. Little, if any, adjustment of the Strategic Plan would be necessary to further align it with the SDGs prior to the completion of the strategy cycle in 2017.

As part of the strategic planning process, each thematic area carried out global-level situation assessments and analyses that resulted in an overarching Programme Division input document (2012). These materials present a rich review of the evidence base, consideration of macro-level trends, unfinished business, strategies and implications, and are another basis for assessing the relevance of the Strategic Plan. These documents anticipate an environment with contracting resources and expanding competition in which strategic choices are needed. In particular, note is made of the need to ensure continuity, make decisions on activities and commitments that may no longer be justified and approach new areas of strategic importance on a selective basis.

There is considerable consistency between the trends and priorities described in the global situation assessment and analyses and the end-of-cycle review.¹⁰ Both documents point to macro-level trends that increasingly affect the lives of children. These include accelerated urbanization and migration, conflict, disaster and climate change, as well as the implications of a demographic transition, which necessitate additional investments for adolescents and youth.

Both exercises voiced strong support for equity as the guiding principle for the new strategy. A consistent theme was the need to utilize UNICEF's wide experience in programming within both development and humanitarian contexts to address underlying vulnerabilities and develop programmatic responses for children and women living in fragile contexts. This emphasis was linked, in part, to strengthening capacities at national and subnational levels to programme for resilience.

These analyses provided a critical knowledge base for evaluating the relevance of the Strategic Plan, as they outlined new areas of strategic importance and emphasized the assessment of new programme areas based on comparative advantage, complementarity and available resources. While such strategic choices were clearly made, no consolidated record was available that described how these strategic choices were weighed and taken. Such a record would likely need to be created for a complete evaluation of relevance.

Key finding 2: The Strategic Plan is relevant to the organizational mandate and international agreements to which UNICEF is bound, including the SDGs.

¹⁰ United Nations Children's Fund, *Report on the end-of-cycle review of the medium-term strategic plan 2006-2013*, UNICEF Executive Board first regular session 2013, Item 5 of the provisional agenda, New York, 7 December 2012, <http://www.unicef.org/about/execboard/files/2013-4-End-of-cycle_review-MTSP-ODS-English.pdf>, accessed 12 August 2015.

As assessed against the UNICEF mandate, other foundational documents and international agreements and accords, the Strategic Plan was found to be highly relevant. For forward-looking goals, there is a high degree of correspondence between the Strategic Plan's core components and the SDGs, both at the level of goals and sub-goals.

3.2 The Strategic Plan as a framework

In addressing evaluability in principle, the evaluability assessment sought to determine the extent to which the Strategic Plan, Theory of Change and results framework present a clear and comprehensive causal chain. However, in examining this issue, it was apparent that the Strategic Plan is both viewed and implemented as an overarching and strategic framework under which programme countries determine the direction and details of implementation. Decisions on programme priorities are based on national priorities, each country's situation analysis and the role of UNICEF within the United Nations Development Assistance Framework (UNDAF). In creating country programme documents, country offices are instructed that the Strategic Plan should guide the descriptions of programme components and outcomes.¹¹ In both the Strategic Plan and the Theory of Change document, UNICEF recognized the balance needed between, on the one hand, sufficient rigor and specificity to provide strategic guidance, and on the other hand, the considerable flexibility to make decisions based on context and led by national governments (*see the adjacent boxed text*).

UNICEF recognition that the Strategic Plan will be implemented flexibly:

"The seven outcomes of the Strategic Plan are sufficiently specific to focus UNICEF on its areas of comparative advantage. At the same time, they are sufficiently broad and flexible enough to enable UNICEF to emphasize cross-sectoral approaches as appropriate and to respond to the specific priorities of individual countries" (Strategic Plan, p. 15).

The 'strategic plan versus framework' discussion has implications for evaluability both in principle and in practice. In this section, this issue is examined as it affects the clarity and comprehensiveness of the causal chain.¹² In Chapter 4, implications for evaluability in practice will be addressed.

For the purposes of the evaluability assessment, we consider a **strategic framework** to be an overarching structure that sets out common goals and themes for the organization and loosely describes a logical sequence leading to those goals. The intention of a strategic framework is to set broad parameters under which more detailed action plans are prepared.

A **strategic plan** has similarities in that it is a document that communicates organizational priorities and agreements around intended outcomes/results. However, a strategic plan would typically go beyond that purpose to focus and direct resources in order to achieve intended

¹¹United Nations Children's Fund, Programme Instruction: Revised Format for Country Programme Documents submitted and approved at the second regular session of the Executive Board 2014, CF/PD/PRO/2014-002, addendum to CF/PD/PRO/2013-005, 2014.

¹² The assessment also drew from a review of 36 strategic plans from United Nations agencies conducted by the Joint Inspection Unit of the United Nations. This review did not produce definitive findings but clarified that the UNICEF Strategic Plan presents the needed characteristics of a strategic plan to either a strong or moderate degree. The findings of this review appear in Annex 5.

results and provide means of gauging progress.¹³ Therefore, a strategic plan would be inherently more prescriptive through the allocation of resources and by way of the axiom ‘what gets measured, gets done’.

The analyses found that the UNICEF Strategic Plan is a corporate-level plan that articulates a results structure from outputs to impact. However, as per UNICEF programme guidelines, the corporate strategic plan is a distal input in comparison to the direct influence of country-stated priorities, analyses of the country context and joint United Nations plans. UNICEF strategy implementation depends on myriad country-by-country selection of outcomes, outputs, implementation strategies and activities. As a result, it is not possible to determine or direct the relative effort that will be accorded to any type of result over another (e.g. demand-generating efforts versus delivery of services); a point that was well-recognized by UNICEF in the planning process. For a decentralized organization like UNICEF, it is not advisable to have a rigidly applied plan.

The evaluability assessment’s analysis sees a strategic framework as an appropriate vehicle for the expression of high-level intent. However, there are implications. The causal chain can be expected to be more consistently applicable at higher levels (impact and outcomes) and far less so at lower levels. At the level of outputs and activities, a rich and diverse set of results can be identified but not easily aggregated. Evaluability of the Strategic Plan will be fundamentally affected by the balance found in strategy implementation across countries and regions. Given the decentralized nature of decision-making in UNICEF, it would be a mistake to evaluate the Strategic Plan as one which is firmly applied across levels or strongly directional. Indeed, the challenge to evaluation will be to identify how the Strategic Plan has guided dialogue and decisions at the country level in a nuanced manner.

Key finding 3: The evaluability assessment found that the UNICEF Strategic Plan serves as an overarching framework.

Key finding 3 is consistent with the majority of interviewees, who viewed the Strategic Plan as providing a common set of high-level results, basic logic sequence and structure around which to discuss programmatic results. This represents a significant challenge to the evaluability of the Strategic Plan in principle as it looks like a plan but is applied like a framework.

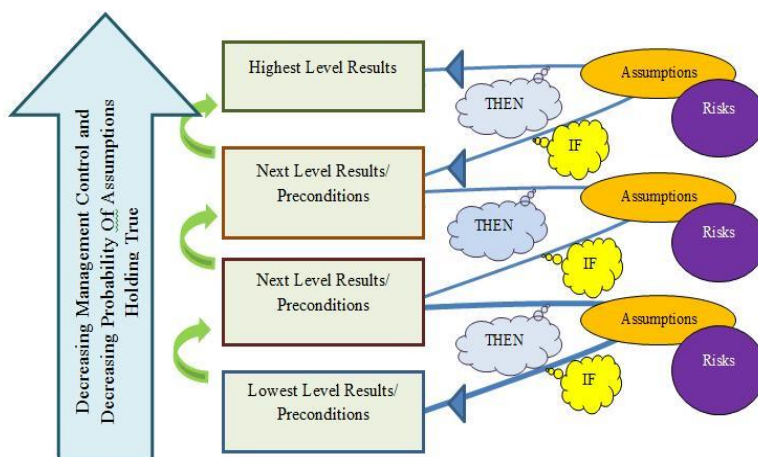
3.3 Clarity of the causal chain

The evaluability assessment terms of reference called for an assessment of the integrity of the causal chain, which was conducted by applying the concepts of ‘plausibility’ and ‘testability’ to the Theory of Change. Plausibility and testability are discerned by examining whether there is a continuous causal chain that connects outputs and outcomes to the final impact of concern. A plausible and testable theory of change also identifies which linkages in the causal chain will be most critical to the success of the strategy, and thus points to potential focus areas for evaluation questions. The Theory of Change was the core document informing this assessment, supplemented with key informant interviews and some additional document review.

¹³ Definitions adapted from: The Balanced Scorecard Institute (www.balancedscorecard.org) and United Nations, *Strategic Planning in the United Nations System*.

As defined in UNICEF internal guidance, a theory of change is a blueprint for building blocks to achieve long-term goals and should represent both how results will be achieved and markers for progress. A theory of change should identify: a) the results sought by a development effort; b) the actions necessary to produce those results; c) the events and conditions likely to affect achievement of results; d) any assumptions about cause and effect linkages; and e) an understanding of the broader context in which the programme operates.¹⁴ Figure 1 provides a depiction of a theory of change with levels and linkages.

Figure 1: Theory of Change logical structure



Source: *Theory of Change in UNICEF Programming. Programme Guidance. Division of Policy and Practice, 2013.*

Outcome to impact level

At the highest level of this framework are impact and outcomes, which the Strategic Plan clearly states are results produced through the shared actions of Member States with support from partners. The mechanisms by which each outcome contributes to realizing the rights of the child are not elaborated in the Theory of Change. In order to visualize the linkages between outcomes and desired impact, it was possible to map each of the outcomes and each of the impact level indicators to the articles of the Convention on the Rights of the Child. Mapped in this way, the underlying connections that relate or associate the outcomes to the desired impact become marginally clearer. In terms of causal links, rather than an 'if-then' statement, the seven outcomes are said to converge to achieve impact.¹⁵ This implies that at the highest levels of the plan, the thinking is not clearly laid out about what is necessary and sufficient to achieve the realization of the rights of every child or in what relative measure.

It should be noted that some sections have developed theories of change for the work of their section. While this is an appropriate place to expound details of a theory of change, the lack of clarity about how each outcome links to the organization's intended impact means that the Strategic Plan Theory of Change cannot be used to make trade-offs or weigh decisions across the outcome areas should the need arise. Since there is no stated theory at this level, no evaluation or measurement scheme can help the organization answer potential questions such as:

- Which of our outcome areas appear/s to have more impact in realizing the rights of the child relative to the others?
- What do we need more of and what can we do with less of?
- Is there a diminishing return on investing further in one outcome area relative to scaling up even more in another?"

¹⁴ United Nations Children's Fund, 'Theory of Change in UNICEF Programming', Programme Guidance Division of Policy and Practice, PowerPoint presentation, 2013.

¹⁵ The UNICEF Strategic Plan 2014-2017, p. 6.

Output to outcome level

At the next level down is a set of six outputs that are similarly framed across outcomes. In the Strategic Plan, this typology is viewed as providing coherence across the results areas and an important underpinning to the organizational Theory of Change.¹⁶ There are pros and cons to the use of standardized outputs as an agency-wide construct to ensure coherence. As noted in the Theory of Change document, they are seen as a unifying factor with the potential to bolster synergies.¹⁷ However, not surprisingly, the aptness of the outputs differs across outcome areas, regions and countries. In some cases, sectors have developed their own technical theories of change. In the Theory of Change, all sectors have added a set of programme areas to further elaborate their priorities, linked implementation strategies relevant to specific outputs and made mention of related activities.

Another important parameter at the level of outputs is that their bases differ as described in Table 5 below.

Table 5: Outputs grouped by varied bases¹⁸

Programming	Crosscutting themes	Supportive action
(1) Demand generation	(4) Humanitarian action	(6) Global and regional programme
(2) Supply of services	(5) Gender equality and rights	
(3) Enabling environment		

The first set of outputs (programming) is consistent with the determinants framework and analyses utilized across country offices to establish a programme mix in individual country settings. Staff across all levels consistently cited these three outputs as an appropriate means of integrating or mainstreaming the Monitoring Results for Equity System (MoRES) into the results framework. Quality, a fourth determining factor of that framework, is not included among the Strategic Plan outputs. Several staff members found this to be an important gap.

Themes that cut across the Strategic Plan comprise the second set of output-level results. Humanitarian action featured prominently in the MTSP end-of-cycle review in terms of the need to anticipate increased frequency and severity of crises as well as the need to integrate humanitarian and development actions. Support for a more integrated approach to the agency's humanitarian and development programming was also voiced throughout staff interviews.

The output/crosscutting theme related to gender equality and rights reiterates two of the three normative principles described in the Strategic Plan. A Gender Action Plan for the agency

¹⁶ Note that these outputs are common for the corporate-level Strategic Plan. Regional and country offices have their own context-specific results statements at both outcome and output levels.

¹⁷ United Nations Children's Fund, Revised Supplementary Programme Note on the Theory of Change for the UNICEF Strategic Plan, 2014-2017, UNICEF Executive Board annual session 2014, 6 May 2014, <http://www.unicef.org/strategicplan/files/2014-CRP_14-Theory_of_Change-7May14-EN.pdf>, accessed 12 August 2015, p. 5.

¹⁸ The complete phrasing of the six outputs appears on page 7 of this report and is not repeated here.

further reinforces gender aspects of the outcomes and further elaborates on logic and cohesion across the identified priorities. The sixth output is related to the UNICEF global and regional programme and its critical role in evidence generation, policy dialogue, advocacy and partnerships.

In principle, the challenge to evaluability does not come from the use of common outputs but from a lack of explicitly stated linkages between results levels. While the Strategic Plan introduces outputs by stating that the “work of UNICEF and its partners leads directly to outputs”,¹⁹ the analyses were not able to discern whether and how outputs lead to outcomes due to the lack of specifically identified linkages. As with the outcome to impact level, there are no ‘if-then’ statements or other explicitly stated cause-and-effect-type relationships between output and outcomes. While some of this information can be inferred from the text, it remains a challenge to evaluability. There are several cases where the descriptions at the output level are concrete enough to represent a testable element of the Theory of Change.²⁰ Even in these cases, although there is no causal chain described, there is sufficient clarity on what types of actions UNICEF will take in order to achieve the output.

Several other challenges were identified, for instance, gaps and omissions, which pose a challenge for evaluability as the logical chain is incomplete. Described below are gaps that were noted between the statement of the output-level result and the corresponding key areas of work (see *Table 6*). *Table 7* identifies challenges to evaluability that are common to output-level results across all outcomes (with exceptions noted).

Finally, in a few instances, analyses found that evidence to substantiate the effectiveness of an activity was lacking. In some cases, an activity is flagged not only because it figures prominently in the Theory of Change but also because it is used as a performance indicator. On a positive note, several outcome areas (i.e. WASH and child protection) acknowledged gaps in the evidence base and identified research priorities as part of Outcome 6 (global and regional programme). Annex 6 includes items with questions on the evidence base that underlie priority activities and performance indicators.

Key finding 4: A significant challenge to evaluability lies in the lack of explicitly stated logical connections of the causal chain from outputs to outcomes and from outcomes to impact. While results statements appear at each level of the Theory of Change, there is very little description of how one level leads to or contributes to another. In other words, while the structure of a results hierarchy is present, the ‘if-then’ statements needed to plausibly link one level to another are largely absent.

More specifically, the seven outcomes are said to converge to achieve impact but no indication is given about what is necessary and sufficient to achieve the realization of the rights of every child or in what relative measure. This lack of clarity at the outcome to impact level means, for example, that the Strategic Plan Theory of Change cannot be used to make trade-offs or weigh decisions across the outcome areas should that question or necessity arise. In addition, the analyses could not discern whether and how outputs lead to outcomes due to the lack of specifically identified linkages: there are no ‘if-then’ type statements or other explicitly stated cause and effect-type relationships.

¹⁹The UNICEF Strategic Plan 2014-2017, p. 9.

²⁰ Examples include: Output 2/HIV/AIDS; Output 2/nutrition; Output 3/WASH.

Table 6: Gaps in the casual chain for specific outcomes

Outcome Area	Output	Gap
Health	Enhanced support for children and caregivers, from pregnancy to adolescence, for improved health behaviours.	Despite their inclusion in the output statement, there is no mention of adolescents.
HIV/AIDS	Enhanced support for children and caregivers for healthy behaviours related to HIV and AIDS and use of relevant services, consistent with the UNAIDS Unified Budget, Results and Accountability Framework (UBRAF).	Not clear how the UBRAF specifically relates to behaviour change and services for the target populations. If the UBRAF sets parameters for UNICEF actions, these are unspecified.
WASH	Increased national capacity to provide access to sustainable safe drinking water and adequate sanitation.	UNICEF's work on water supply is described as shifting, with diminished emphasis on direct service delivery and increased emphasis on capacity development. No rationale is given for this change in programming shift, nor is it explained which actors will fill the gap in service provision.
Nutrition	Enhanced support for children, caregivers and communities for improved nutrition and care practices.	With the exception of one reference to UNICEF, the key areas of work make no mention of actors, thereby clouding accountability for the achievement of the result and making it difficult to plausibly link or test the association between these activities and the output.
Education	Increased capacity of governments and partners, as duty bearers, to identify and respond to key human rights and gender equality dimensions of school readiness and performance.	The output statement includes 'school readiness', which is not explicitly addressed in the narrative.

Table 7: Gaps in the causal chain common across outcomes

Output	Issue	Observation
1	Not stated as a result	While other outputs focus on variants of capacity strengthening, Output 1 is framed as “enhanced support”, more accurately considered as a process. Child protection is an exception in that both “enhanced support” and “increased capacity” are used.
3	Covers an overly broad range of results and actions	Output 3 represents a multiplicity of results (i.e. political commitment, accountability and national capacity) and actions (i.e. legislate, plan and budget). No outcome area covers all of these concepts nor are they defined or weighed in regards to effect on scaling up. Indeed, the output phrasing is inconsistent with other UNICEF guidance, which advises that outputs should not combine many different types of change. ²¹
4	Little focus on national capacities	This output is phrased as “increased country capacity and delivery of services for humanitarian situations” but deals most directly with either UNICEF’s role in coordination mechanisms or approaches for direct service delivery. Even where capacity strengthening is mentioned (HIV/AIDS, WASH, nutrition, child protection), there is very little description provided (often one sentence). Social inclusion is a notable exception with its more detailed description of capacity building efforts. It should be noted that elements of preparedness planning do appear under Output 3 (i.e. health, education).
5	Inadequate attention to human rights	While this output refers both to human rights and gender equality, it was found to primarily address capacities related to gender. Several outcomes made little reference to rights (health, HIV/AIDS, WASH, nutrition).
6	Lacks reference to/definition of capacity as well as how capacity will be enhanced	While the Strategic Plan stated output deals specifically with “an enhanced enabling environment”, all seven outcomes have alternatively used language related to global and regional capacity. However, Output 6 descriptions lack specific reference to and definition of capacity, as well as how capacity will be enhanced. One is left to infer that the result (enhanced capacity) is equated with the conduct of the noted implementation strategies.

3.4 Key assumptions, risks and mitigation strategies

The Theory of Change includes a description of key assumptions, risks and mitigation strategies and notes that the UNICEF approach was informed by key sources, including guidance from its Executive Board, dialogue with United Nations sister agencies, and consideration of the organization’s needs.

Assumptions are the unstated or implied beliefs or understandings (conditions that must hold true) that pertain to the transitions between levels (i.e. activities to outputs, outputs to outcomes, outcomes to impact). Assumptions should describe events or conditions that are highly likely to occur but with at least some small probability that the assumption will not hold.

²¹ UNICEF, ‘Results planning for 2014 and the new coding structure’, PowerPoint presentation.

The Theory of Change presents several sets of assumptions that differ according to the level of the results framework and may be either generally applicable across the level (e.g. a set of assumptions pertaining to all outcomes) or may be specific to a given outcome area. A summary of these distinct characteristics and crosscutting assumptions, as well as outcome-specific assumptions, appears in Annex 7.

The analyses found that the assumptions at the level of outcomes to impact are framed consistently. Per the definition, outcome to impact level assumptions are implied beliefs or understandings with a high likelihood of occurring. At the level of outputs to outcomes, however, several assumptions are phrased in a manner that is counter to the UNICEF definition of assumptions as implied beliefs or understanding with some small probability of not holding true. For example: resources for the issues addressed by the outcomes do not disappear entirely; political support for the issues covered by the outcomes does not vanish; and the Strategic Plan does not assume continued growth in resources, but does rely on a belief that they will not completely dry up.

Assumptions about strategic partnerships also seem to be covered only lightly. Indeed, much of the change sought in the Strategic Plan depends more on the performance of partners than on the performance of UNICEF. Given the Strategic Plan's wide recognition of the importance of partnerships, more specific assumptions regarding their roles would have been welcomed.

Perhaps most notable in regards to the assumptions is the degree of repetition. While the Theory of Change notes that certain of the assumptions appear at multiple levels,²² the degree of replication suggests a cookie cutter approach to the development of the assumptions, as opposed to sector-specific reflections on the pathways to change, including definition of critical assumptions.²³ This is the first attempt at creating a theory of change for the Strategic Plan. It is possible that the MTR will throw better light on the reliability of the assumptions and how they have been managed. This needs to be taken into account in an evaluation.

²² Excerpt from Revised Supplementary Programme Note on the Theory of Change for the UNICEF Strategic Plan, 2014-2017, p. 11: "... an assumption may apply to more than one level in the causal pathway. For example, a massive increase in humanitarian crises (e.g., due to the outbreak of a novel pathogen) would affect the ability of outputs to lead to an outcome, but it would also jeopardize the transition from outcome to impact. This is because even if changes have occurred in institutions and/or if individuals have changed their behaviours, these shifts may be insufficient to improve the lives of a child confronted with a devastating pandemic."

²³ For example, all seven outcomes articulate the assumptions: "Communities are open to receiving information, behaviour change messages and other efforts" (Output 1); "A minimum level of capacity exists such that capacity development efforts do not falter" (Output 2); "The global system of coordination and partnership does not collapse and be replaced with increased competition and rivalry" (Output 6).

4. Evaluability of the UNICEF Strategic Plan in practice

This chapter concerns evaluability in practice and seeks to answer the question: how well can the Strategic Plan be assessed in terms of the measurement approach, tools and indicators used to reflect performance and results achieved? The complete set of questions to analyse evaluability using this lens appears in the methods section in Chapter 2 of this report. This chapter addresses the topics listed under Objectives 2 and 3 in the evaluability matrix and Objective 2 in the terms of reference (both found in the inception report in Annex 1). A summary of the key findings related to evaluability in practice appears in a box at the end of the chapter. The implication of findings for the Strategic Plan overall is also highlighted.

This chapter also continues the discussion of the Strategic Plan as an actionable plan, *per se*, versus its application as an overarching framework.

4.1 Indicators, baselines and targets

The extensive use of performance indicators is the primary way progress measurement and monitoring will be carried out for the Strategic Plan. Baseline values of these indicators are crucial for performance measurement since they provide the starting point from which progress will be gauged. Likewise, the use of targets provides a clear statement of what success looks like, permitting accountability to be managed and progress to be more easily interpreted.

There are a total of 165 Executive Board-approved programmatic indicators in the results framework. These cover the impact level, all seven outcomes and their respective outputs.²⁴ However, many of the indicators were either constructed as compound indicators²⁵ or required reporting in a disaggregated manner (e.g. by gender for the purposes of equity monitoring). This means that the actual number of indicators intended to be measured was significantly greater than 165. A little more than half of the programmatic indicators in the Strategic Plan will be generated through country office reporting. In 2014, a set of 35 supplemental indicators was developed. These were intended to make measurement of the causal chain, from outputs to outcomes, more robust. Given the heavy reporting burden on country offices, a decision was made in late 2014 not to implement these supplemental indicators. Even without the list of supplemental indicators, the results framework comprehensively covers impact, all outcomes and every output, with multiple indicators for each level.

The in-depth analysis of indicators here includes all impact and outcome level indicators, and a sample of about 25 per cent (26 of 105) output indicators, as described in the methods

²⁴ This does not include 60 indicators of organizational efficiency and effectiveness that are published in the results framework.

²⁵ A compound indicator requires the separate reporting of sub-components as in the case of this health indicator: “P1.5 Countries with at least 90% coverage at national level, at least 80% coverage in every district or equivalent administrative unit for children <1 receiving (a) measles-containing vaccine, (b) DTP-containing vaccine.”

section. The inception report describes the selection of indicators for in-depth analysis and Annex 8 contains the indicator-by-indicator analysis for the entire sample.

Table 8 and Table 9 summarize the status of baselines and targets in the sample of indicators. Among the 86 indicators, seven at the outcome level and five at the output level lacked baselines. Where baselines were provided, the results framework did not indicate the source of the baseline measures. These were presumed to be the same as those required for ongoing indicator measurement. This should be verified given that if baselines and ongoing measurement are not based on the same source, there is the potential that documented change in the indicator could be misinterpreted. It should also be noted that among all of the indicators, 40 (approximately 25 per cent) are to be disaggregated by sex, urban vs. rural residence, and/or wealth quintile, where possible. The purpose of disaggregation is to provide insight into performance gaps between relatively more and less advantaged groups (e.g. urban vs. rural residents). However, only five disaggregated baselines were provided. The aggregated baselines provided will only be informative if aggregated measures of the indicators are reported along with disaggregated values.

Performance targets were assessed using a three-category rubric: 1) red = no description of how the target was set or no target; 2) yellow = target based on expert opinion with no further explanation; and 3) green = target set through a transparent process and/or aligned with global or regional commitment. Six of the impact indicators lacked targets. Among the outcome and output indicators reviewed, six also fell into the category of no target or no description. Twenty-four outcome and output targets were set based on expert opinion from technical leads at UNICEF. The majority (40) were set either using a clearly articulated and transparent process or were aligned with some form of global or regional commitment for the relevant sector.

Based on volume alone, there are more than adequate numbers of indicators spread across the impact, outcome and output levels to inform performance management. The indicators generally have a baseline reported based on some national aggregate figure of performance, even though roughly one in four indicators is meant to be disaggregated by sex or some proxy for relative poverty or disadvantaged status. Most indicators have targets set, the majority of them through a transparent process or in alignment with international or partner agreements.

Table 8: Assessment of baseline in place

Level	Baseline in place?	
	Yes	No
Impact	16	0
Outcome	37	7
Output	21	5

Table 9: Assessment of target identified

Level	Ratings		
	Target set through transparent process and/or aligned with global or regional commitment	Target based on expert opinion with no further explanation	No description of how target was set or no target
Impact	7	3	6
Outcome	28	12	4
Output	12	12	2

4.2 Quality of the indicators and implications

As described in the methods section in Chapter 2, the four aspects of indicator quality were examined. These are summarized in Tables 10-14.

Table 10: Assessment of indicator quality: Is indicator defined?

Level	Ratings		
	All indicator components clearly and objectively defined; little or no room for subjectivity	All indicator components clearly defined with some elements of subjectivity remaining	Not defined, or defined with one or more elements of a highly subjective manner
Impact	16	0	0
Outcome	41	3	0
Output	18	6	2

Table 11: Assessment of indicator quality: Can indicator be reliably measured?

Level	Ratings		
	Clear or standard indicator easily measured using well-established practice	Indicator with weak specification or source; or dependent on subjective assessment (e.g. by country office staff person)	Poorly constructed; framed in immeasurable manner; or measurability unclear even from specifications provided
Impact	13	3	0
Outcome	39	4	1
Output	15	9	2

Table 12: Assessment of indicator quality: Is indicator specific?

Level	Ratings		
	Indicator matches type of result anticipated (impact, outcome, output)	Possible relationship to type of results anticipated (impact, outcome, output)	Mismatch between type of results expected and indicator (impact, outcome, output)
Impact	15	0	1

Outcome	42	2	0
Output	15	10	1

Table 13: Assessment of indicator quality: Is indicator sensitive to change?

Level	Ratings		
	Yes	No	Unknown
Impact	4	7	5
Outcome	36	7	1
Output	23	0	3

Table 14: Data quality checks

Level	Ratings		
	Yes	No	Unknown
Impact	13	1	2
Outcome	37	6	1
Output	16	10	0

For the vast majority of the indicators, definitions and computation methods for reliable measurement were provided in the detailed indicator sheets. The majority of indicators without proper definitions were at the output level, with 8 of the 26 output-level indicators sampled found to be lacking in elements of good definition.²⁶ In addition to indicators lacking good definitions (which, logically, will not be reliably measurable), eight other indicators at the impact, outcome and output levels were found to lack good and reliable measurability. The computation methods required for indicator calculation ranged in complexity. In some cases they required several data elements or multiple steps in calculation.²⁷

The specificity of indicators was somewhat more consistent. At the impact level, there was one indicator at the level of an output (i.e. a measure of execution) not at the level of a major or sustainable result in the target population (i.e. impact among children central to the realization of their rights). At the level of the outcomes, two indicators were not strongly related to the outcome. At the output level, 11 indicators did not correspond to output-level results; these tended to relate to results that should be considered outcomes.

The two other aspects of indicator quality were: sensitivity to change during the strategy period, and whether there are clear procedures in place for quality assurance of the indicators reported. At the impact level, sensitivity to change during the strategy period was infrequent, with only four indicators seeming to be clearly sensitive to change.²⁸ At the outcome and output levels, sensitivity to change was more likely. Fifty-nine out of 70 outcome and output indicators were felt to be sensitive to change during the 2014-2017 period.²⁹

²⁶ Across the indicators, two were found to not be defined, or were defined with one or more elements of a highly subjective manner; nine were found to have some elements of subjectivity remaining in the definition; and the remainder were clearly and objectively defined with little or no room for subjectivity.

²⁷ For example, P7.4, as described in Indicator Sheets for UNICEF Strategic Plan, 2014-2017 Outcome 7: Social Inclusion, p. 25.

²⁸ Of the remaining indicators, five were of questionable sensitivity to change, and the remaining indicators were unlikely to change materially during the strategy period.

²⁹ Of the eleven indicators that were either questionable or unlikely to change in the strategy period,

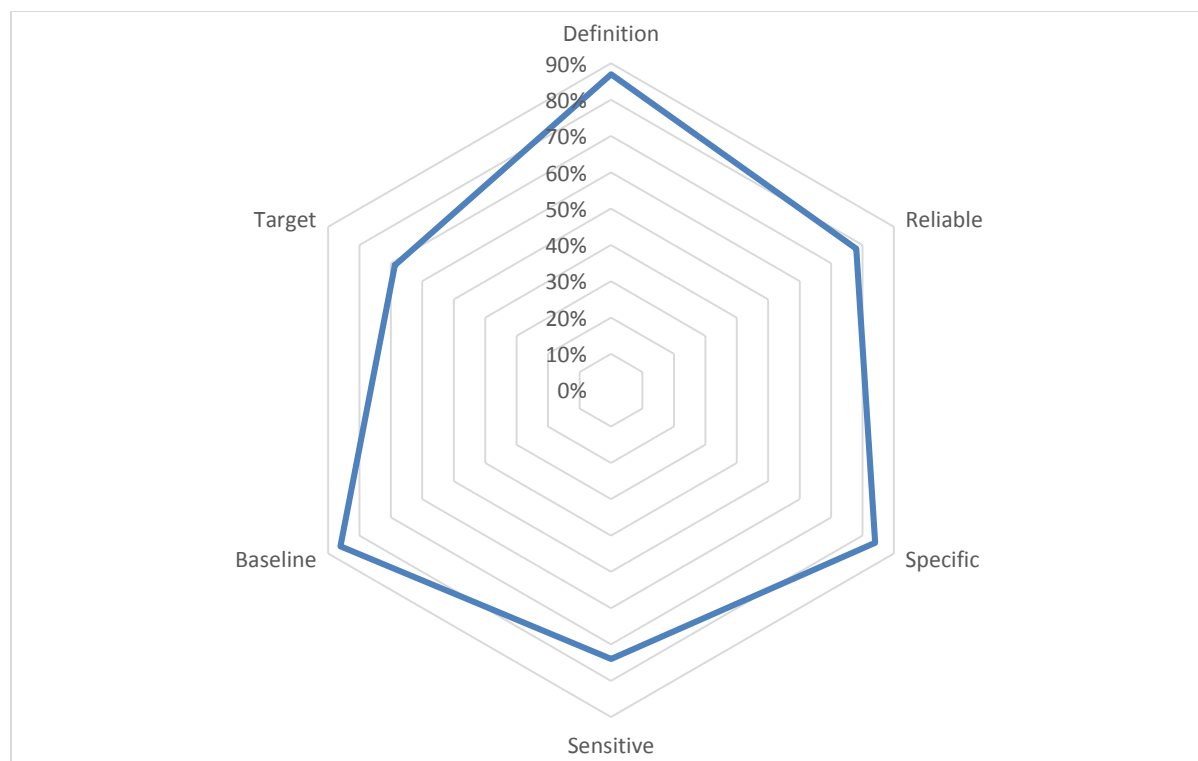
With respect to quality assurance measures for indicator data, some patterns emerged. At the impact level, almost all indicators are the product of rigorous and well-established methods that include quality checks on data and computation of indicators. While two of three impact indicators for social inclusion are new and do not have clear quality checks, only one (indicator 7a) is likely to be problematic with regard to uncorrected or unchecked quality. Within the outcome areas, 10 of 26 output level indicators lack known quality checking mechanisms. For data reported from country offices where the assessment of country office staff forms the basis of indicator reporting, there may be a lack of well-established data quality checks in place. This potential weakness in data quality has been acknowledged at the Headquarters level and is being addressed to some extent with the introduction of logic checks in the new version of the RAM system (see below). This situation contrasts with the narrative reports of progress submitted into the RAM system, for which there are several layers of editorial review before inclusion in annual report documents.

There have already been lessons learned regarding the quality of the performance data collected from country offices. In early 2014, an initial effort to collect indicator data from country offices via the country office annual reporting system resulted in data of variable quality, based primarily on lack of consistency. To ensure greater consistency, the seven outcome areas developed detailed indicator sheets. For each indicator, these sheets contained: an indicator definition and calculation method (e.g. definition, method of computation, whether the indicator is cumulative or specific, universe of countries to report/report on); a source for the data and responsible unit; and frequency of updating, required disaggregation and the target-setting process. The next cycle of annual reporting, which will include re-creation of the baseline, was launched in late 2014.

In sum, it should be noted that indicator quality ranged widely when rated by these rubrics. Figure 2 provides the overall percentage of indicators meeting the highest standard of the rubric for each criteria. Criteria including indicator definition, existence of a baseline and specificity in relation in the intended result scored well (87 per cent, 86 per cent and 84 per cent, respectively) across the set of indicators examined. Reliability and sensitivity to change fared slightly less well with the average percentage of indicators achieving the highest standards at 78 per cent and 74 per cent, respectively. The availability of transparently set targets was the weakest area, with only 55 per cent of indicators judged to have met the highest standard. In other words, more often than not, indicators had the necessary attributes of being well defined, with a baseline, reliably measurable, specific and sensitive to change.

eight were outcome indicators and three were at the output level.

Figure 2: Criteria-based assessment of indicator quality, 86 indicators in results framework



The quality of indicators showed differences across levels with output indicators far more likely to be rated poorly against rubrics. The two exceptions to this observation, that is, where indicators were judged poorly at all levels, related to transparently set targets and sensitivity to change. Indeed, one potential challenge to evaluability is related to the impact-level measures, as many of those indicators are either not sensitive or are of questionable sensitivity to change within the 2014-2017 timeframe. Impact indicators that are not measurable and that change within the four year period will complicate judgment of progress in the Strategic Plan at this level, irrespective of how well specified the indicators may be and regardless of how much progress has been made and measured at the outcome level. This challenge is not limited to UNICEF because the chosen impact-level measures are largely international standards. Similarly, 8 out of 55 outcome indicators were also thought not to be sensitive to change. Given the time-limited strategy period (four years), expectations of measurable change, for example as held by the Executive Board, should be tempered.

Other areas of concern include the 11 indicators with weak or absent indicator definitions and the 16 indicators with concerns over reliability in measurement. Without common definition and clear, objective guidance on the calculation of the indicators, there is naturally an opportunity to introduce a greater degree of subjectivity, errors and inconsistencies in measurement. The majority of indicators with these deficiencies were at the output level. Given the ratings on the reliability of measurement for output indicators, care should be exercised in their interpretation. The utilization of disaggregated data in monitoring performance is unclear (e.g. whether baselines will be available in disaggregated form), yet seemingly an important element of the equity agenda. Finally, expectations should be

tempered for measureable change during the strategy period at the impact level.

In general, the evaluability assessment found that the indicators examined performed well against the criteria, including in terms of indicator definition, the existence of a baseline and specificity to the intended result. The availability of transparently set targets was the criterion with the poorest overall scores. Some deficiencies were also noted regarding those indicators meant to be reported in a disaggregated manner. Indicators with quality concerns exist predominantly at the output level.

Key finding 5: The assessment found that the indicators examined performed well against the criteria used for this exercise, including in terms of indicator definition, the existence of a baseline and specificity to the intended result. However, a few performance indicators as found in the results framework are weaker and in some cases, may impede evaluation through poor quality or limited interpretability.

4.3 Evaluability of normative principles and crosscutting issues

Three normative principles, namely human rights, gender equality and environmental sustainability, provide an underlying basis for UNICEF action. Normative principles are characterized as universal (i.e. applying equally to all people in all countries); based in law, internationally-agreed development goals and treaties and national laws and commitments; relevant to all of UNICEF's cooperation with government and other partners; and differing from priorities and goals that depend on setting and context.³⁰ Accordingly, these principles are clearly articulated throughout the Strategic Plan and the Theory of Change, as well as in a stated intent to improve monitoring and reporting of results related to the normative principles.³¹ Issues pertaining to the evaluability of these normative principles, crosscutting themes (i.e. gender equality and humanitarian action) and equity are considered here.

Equity

The Strategic Plan articulates the UNICEF focus on equity as follows: "Equity means that all children have an opportunity to survive, develop and reach their full potential, without discrimination, bias or favouritism." Although the implications of evaluability in practice are multi-fold, the evaluability assessment findings centre most directly on the need for explicit equity measurement approaches and the dichotomy inherent in aggregating results for a corporate-level plan with data from highly situation-specific programming approaches.

The analysis found that throughout the Strategic Plan, outcome indicators are framed in terms of the numbers of countries that have met certain conditions. For example: "Countries with at least 80 per cent of children aged 0-59 months with diarrhoea receiving oral rehydration salts"; or "Countries in which more than 50 per cent of primary and secondary schools have WASH facilities meeting national standards." In other words, the indicators call for a single, national number, without respect to reaching targets in sub-groups of interest (e.g. girls, the most disadvantaged, or children in the poorest wealth quintile).

Thus it is possible for a country to meet the target threshold and to be counted as a success

³⁰ *Programme, Policy and Procedure Manual*, Section 2.5, Normative Principles.

³¹ The UNICEF Strategic Plan 2014-2017, pp. 11, 17.

without achieving targets in subgroups of interest. The Theory of Change, in fact, calls out the risks of relying on national averages to measure equity: “It is increasingly clear that national average figures tell only partial stories.”³²

With this in mind, the Strategic Plan acknowledges the need for disaggregated sub-national data, citing the development of capacity to generate it as a priority: “A growing priority is to strengthen national and subnational collection, analysis and use of data, particularly on children and women. This will include more systematic use of disaggregation to enhance equity-focused design and implementation of policies and programmes.”³³ Indeed, in many cases where disaggregated indicator data are preferred, the results framework acknowledges they may not be available, often calling for disaggregation only “where feasible.” Further, UNICEF made a commitment to developing a comprehensive measure of equity that is applicable in varied contexts while maintaining cross-country comparability to the extent possible.³⁴

The risks to good measurement stem not only from uncertainties about how comprehensively and reliably disaggregated data will be available. Part of the risk also stems from differing understandings among staff of the organization’s approach to measuring equity-related programming. Several of those interviewed pointed out that the question of the organization’s approach remains largely unresolved. Rather than an organization-wide approach, sections and country programmes have taken different approaches to the challenge of making the goal of equity-focused programming more explicit in their measurement frameworks. For example, the education section has invested heavily in developing guidance materials for more consistency and reliability in understanding and reporting on equity-related programming.

When equity is invoked as a desired result, goal or objective, there is a need to be explicit about a vision of success. Given the centrality of equity to the Strategic Plan, UNICEF should be more explicit about expected results. A well-constructed theory of change will support the organization in not only measuring progress in this area but also better defining roles and responsibilities at the various levels of the organization.

Given the centrality of equity to the Strategic Plan, one would expect to find a cogent discussion of the approaches that UNICEF is employing to measure progress in this area with roles and responsibilities outlined at Headquarters, regional office and country office levels. Currently, it is not evident how equity would be evaluated. While different parts of UNICEF appear to have implicit ideas about equity, few of these are explicit enough to permit measurement and evaluation.

Key finding 6: The entire Strategic Plan is framed with an equity lens, yet the measurement of equity in practice is weak in the Strategic Plan and is further hampered by the questionable availability of data to measure progress in this area.

³² Revised Supplementary Programme Note on the Theory of Change for the UNICEF Strategic Plan, 2014-2017, p. 8.

³³ The UNICEF Strategic Plan 2014-2017, p. 12.

³⁴ Ibid., p. 17.

Human rights and gender equality

Human rights and gender equality are integrated throughout the Theory of Change document as foundational principles that underlie all UNICEF programming. Gender equality has also been identified as a crosscutting theme for the Strategic Plan and both principles are given structure as an output (i.e. “Increased capacity of governments and partners, as duty bearers, to identify and respond to specific challenges related to the protection and promotion of the rights of children and gender equality”).

Several challenges were noted regarding the evaluability of the principle of human rights. First, realizing the rights of the child is the impact of the entire Strategic Plan and rights-based approaches underpin UNICEF programming (e.g. as one senior manager said “the entire plan is about rights”). Placing rights as part of a separate output brings it visibility. However, it may also contribute to a perception that rights are addressed under this output alone as opposed to cutting across all of UNICEF’s work. Placement of rights as part of an output affords visibility but may lead to a mistaken impression that this is the ‘rights’ output in an isolated manner. To be evaluated fully, rights-based approaches will need to be conceptualized as integral to all outputs and outcomes (e.g. support for rights holders in Output 1). There is an underlying dichotomy that needs to be resolved regarding the treatment of rights as a programming approach (as per the human rights-based approach to programming) versus an output-level result related to increased capacity of duty bearers.

The measurement of progress in the area of human rights is not well elaborated. Across the Strategic Plan, only two indicators deal explicitly with the realization of rights (P5.e.2 and P7.e.1). Moreover, staff members view the rights-related performance measures (as attached to Output 5) as weak. Some additional information may be drawn from the country office annual reports, which include a brief description (350 words) of the human rights-based approach to programming.

Gender equality is also a foundational principle for UNICEF. Gender equality appears as a crosscutting theme and is also encompassed within an output-level result. As in the discussion on rights above, placement of gender equality as part of an output gives it greater visibility but may lead to a mistaken impression that gender-focused activities are limited to this output. Efforts to evaluate the Strategic Plan will necessitate a crosscutting examination of gender-focused programming. Limiting the evaluation of gender-focused actions to this single output would be a severe oversight on the part of evaluators.

The gender equality focus of the Strategic Plan is also supported and advanced by the Gender Action Plan. The Gender Action Plan elaborates the gender dimensions of results across the seven outcome areas and presents a programme framework that brings greater cohesion to gender-focused actions, including underlying barriers and bottlenecks. For the purposes of evaluation, the Gender Action Plan should be viewed as a critical accompaniment to the Strategic Plan.

For gender equity, indicators for Output 5 make progress measurement clearer across the strategy, compared with those at the outcome level. A further benefit to evaluability of this crosscutting priority is the Gender Equality Marker. The Gender Equality Marker is a function in UNICEF’s internal programme monitoring that ‘flags’ or identifies activities with significant programme focus and expenditures related to gender equality. In addition, every year, all country offices prepare a brief summary (350 words) to highlight their activities in the area of gender equality.

Environment sustainability

Environmental sustainability is a normative principle that does not appear as a crosscutting priority or distinct output within the Strategic Plan. Accordingly, the relative weight given to environmental sustainability in any evaluation should be far less than that assigned to either rights-based programming or gender equality. An issue for evaluability is the relatively recent appearance of environmental sustainability on the child rights agenda. UNICEF has recognized the need to strengthen its approach to environmental sustainability and is developing a long-term approach with a view to contributing to: a) reducing the negative impacts of climate change and environmental degradation on children's rights; and b) identifying and capitalizing on opportunities arising from the global and national attention to climate change and environmental degradation for children.

Given the recent focus, not surprisingly, environmental sustainability is not mentioned in the impact, outcomes or output narrative in the Theory of Change or results framework. It is virtually unrepresented in the programmatic indicators.³⁵ However, the draft research agenda for the period 2014-2017 identifies several priority questions related to environmental sustainability.³⁶ As with the other normative principles, country offices prepare an annual narrative summary (350 words) to highlight their activities in the area of environmental sustainability.

Humanitarian action

For the humanitarian area of work, there is wide recognition that changes are being put into action, some of which preceded the Strategic Plan. The Strategic Plan seeks to move UNICEF further toward greater continuity of effort across development and humanitarian work and brings a clearer focus on resilience. As with gender equality, an immediately apparent challenge to evaluability would be that humanitarian action is construed narrowly as one among six outputs rather than as a set of actions being undertaken across the organization (e.g. changes in management structures).

Evaluability in practice may also be challenged by recent changes, including the revision of the Core Commitments for Children in Humanitarian Action (CCCs) with the inclusion of commitments and benchmarks based on existing international standards. These commitments and corresponding benchmarks serve to guide country offices on indicator selection and creation, where needed. The use of these benchmarks should facilitate a more structured approach wherein agreements or memoranda of understanding are better able to link resources with results. The utilization of these benchmarks should be closely monitored and subject to evaluation.

In terms of centralized reporting, when completing the RAM, country offices undertaking a humanitarian response will be asked to complete the related strategic plan performance indicators on an outcome-by-outcome basis. Reporting against the CCC benchmarks and indicators will be conducted through a separate component of the country office reporting system. Although the results framework, humanitarian action markers/tag, and revision of the Programme Information Database codes better reflect the integration of systems, a more consolidated approach would also build the CCC benchmarks and indicators into the RAM

³⁵ Indicator P7.c.2 is the only indicator that explicitly mentions environmental sustainability.

³⁶ United Nations Children's Fund, *A Framework for Research 2014-17*, 2014.

system.

An additional challenge to evaluability may arise in variable agency-wide understanding and uptake of preparedness and risk-informed programming. For example, there is a widely recognized need to develop better understanding, definition and measurement of resilience. See more on this challenge in the boxed text.

Perhaps most notable is the humanitarian action theme, which serves as the basis of new programming aimed at greater integration of development and humanitarian actions and as a significant proportion of UNICEF expenditures in recent years (e.g. reportedly around 50 per cent).

Key finding 7: Crosscutting themes pose challenges for evaluability.

Recent evaluation noted challenges to evaluability of preparedness actions both in principle and in practice

A recent evaluation concluded that efforts to assess the effectiveness of emergency preparedness activities and impact-level results were hindered by inconsistent and/or limited data collection and analysis. To illustrate, the agency-wide Early Warning Early Action self-assessment tool that country offices use to identify risk of emergencies and readiness to respond and plan preparedness activities was found to be used inconsistently and with little common understanding among staff, resulting in significant variation in risk assessment practices.

The evaluation also concluded that there is a limited evidence base to foster understanding of how preparedness leads to better response, what works best and how to improve and innovate. The research priorities list (discussed in Chapter 5 below) may help to answer these questions. Topics include costing resilience programming and measuring the impact of integrated/inter-sectoral approach in humanitarian settings.

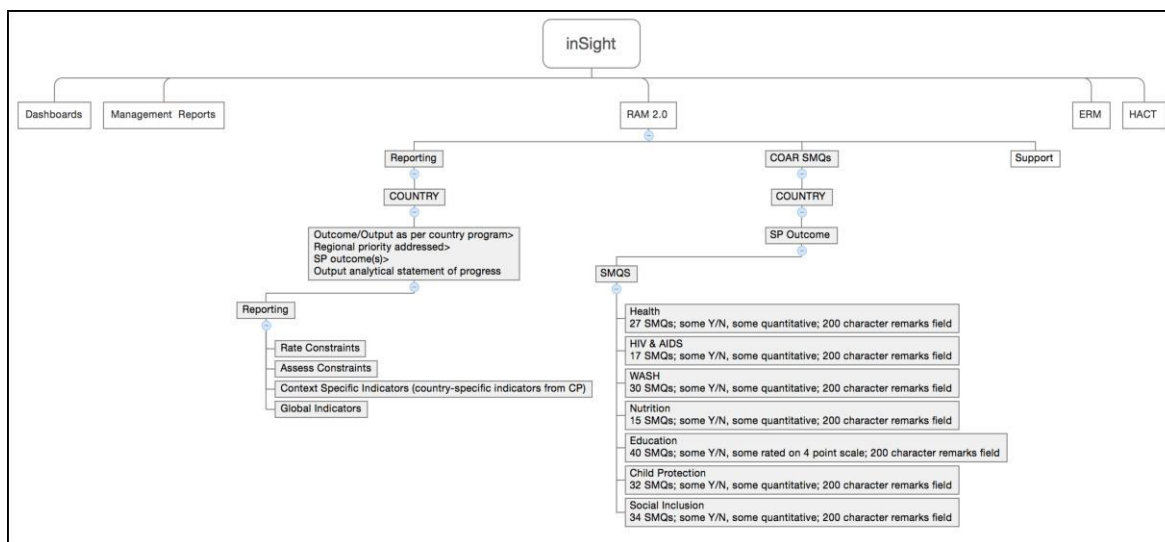
4.4 Data generating and reporting systems

Measurability in practice means not only that the indicators have been well specified and are valid, but that data sources and the systems to generate and report on indicators are also established. Aside from the impact indicators, for which sources were not specified in the results framework (but which derive primarily from well-established sources), data sources have been identified for all indicators. As seen in the results framework, the main sources of information are the country offices, via their annual reporting cycles. The Global Database maintained by the Data and Analytics Section and humanitarian performance monitoring systems are also significant sources of data. Partner agencies are an additional source of indicators. Many of the HIV and AIDS indicators, for example, are intended to come from UNAIDS, while some of the education indicators are supplied by the UNESCO Institute for Statistics, which has the United Nations mandate for collecting statistics on education from governments.

As noted, there are some quality checks in place for the SMQs across the country office annual reporting system. This is the part of the overall measurement and reporting system in which the greatest amount of variation may be expected in terms of quality and consistency (as summarized on the next page). Therefore, efforts have been made to shore it up through built-in quality checks (e.g. encoding of logical skip patterns in online reporting systems) and desk review of data by the Headquarters unit responsible for compiling reports. It should also be noted that there is a well-elaborated system to update indicator estimates for more than 100 outcome and impact indicators, and to assure and adjust for quality concerns in these indicators (i.e. the Country Reporting on Indicators for the Goal or 'CRING'). Both of these systems are discussed.

Considerable effort has been put into data generating systems with clearly defined roles and responsibilities. The primary conduit for reporting on strategic plan indicators is through the ICON country portal and via the country office reporting system. The RAM 2.0 has two main components: a) country office reporting against defined outcomes and outputs in accordance with their CPD resources and results framework and narrative summaries of progress; and b) the Country Office Annual Report (COAR), which includes the SMQs needed to generate the majority of indicators from the corporate-level results framework. This system is diagrammed in Figure 3. As the name indicates, the periodicity of this system is yearly, though certain sections of the reporting system can be accessed off cycle. Information from these sources is ultimately synthesized into reports such as the Executive Director's Annual Report.³⁷

Figure 3: Structure of country reporting system on performance indicators



At the time that the evaluability assessment was conducted, it was not possible to fully analyse the functioning of this system in practice. An early 2014 launch was followed by extensive revisions and the system was re-launched in November 2014. It had therefore not yet completed a full iteration with country offices receiving SMQs, responding to them, and the baseline data being compiled, quality checked, analysed and interpreted.

³⁷ In one exceptional case (i.e. education), results framework performance indicators are compiled, analysed and presented in thematic, regional- and country-specific strategic plan education profiles that support programming at country level as well as dialogue between Headquarters, regional offices and country offices.

In lieu of observation in practice, interview data from Headquarters and regional and country offices provided insight into what was anticipated for this system. Some of the higher-level stakeholders expressed satisfaction with the system and confidence that it will function well. They pointed to the significant amount of work that has gone into the detailed indicator sheets to ensure consistency in responses to the SMQs; the substantial amounts of internal communications and training at the Headquarters level about the measurement of the Strategic Plan; the overall reduction in the number of SMQs to reduce the reporting burden on country offices (a revised approach so that country offices answer only those SMQs relevant to the context); skip-patterns coded into the online reporting system that will remove unnecessary/irrelevant questions from response options available to those responsible for answering the SMQs; and support that will be available during reporting season.

Other stakeholders raised issues or voiced concerns. A complete cycle of reporting will be needed to determine whether common complaints have been remedied in part or in full. Chief among them is the fact that even with skip patterns, the reporting burden may remain heavy. Concerns were voiced about uneven and often inadequate capacity at the country office level for reporting and the likelihood that the quality of measurement via this system could be affected. Recognizing that the COARs can produce data of varying quality, some Headquarters outcome teams use a subsequent verification process with outcome leaders in the country offices.

The relevance of the SMQs to the country programmes was another concern raised frequently in interviews with regional and country office staff. There was a considerable degree of displeasure at regional and country office levels that the reporting (through 2013) required almost all countries to answer almost all questions regardless of relevance to the country programme of cooperation. It will be important to gauge if the revamped COAR/SMQ system, with its more limited burden on country offices, is perceived as making a difference. The reliability of measurement was also a concern, with country office respondents acknowledging that confusion over the instructions and language used can mean that within one country office, one person might answer 'yes' and another person might answer 'no' to the same question.³⁸

A further concern pertains to the COARs/SMQs, which will generate the majority of output-level indicators for the corporate Strategic Plan. The issue deals with the variation in the language used in the SMQs regarding UNICEF involvement in or contribution to the programme area of interest. Some outcome areas have consistently phrased their SMQs to report on progress only as a result of UNICEF direct support or only if the progress is relevant in reference to the country programme; other outcome areas may refer to progress in UNICEF-targeted areas or in UNICEF-supported programmes; and some outcome areas make no reference to UNICEF whatsoever. This degree of variation in language that addresses the specific contributions of UNICEF may create confusion for those completing the SMQs in country offices, as well as for those interpreting the responses.

The Global Database is also an important source for plan indicators. The Data and Analytics Section, part of the Division of Data, Research and Policy, maintains the Global Database, which is cited as the source for 37 of the 150 outcome and output indicators. The data for

³⁸ Again, variation by sector was apparent with the education section developing specific guidance and criteria for four point numeric scales to minimize subjectivity of answers and maximize comparability across countries to aggregate global numbers.

these indicators are drawn from a variety of routine and administrative sources, as well as high quality household surveys, including UNICEF's Multiple Indicator Cluster Survey and the Demographic and Health Survey. The Global Database is resourced to carry out extensive data compilation, quality assurance and analysis. One of the main mechanisms for updating the Global Database is the annual round for CRING. This mechanism requests that country offices provide updated information on 117 indicators, including impact and outcome indicators.³⁹

This reporting, in itself, represents a formidable task – both for those reporting on updated indicators and sources and for those at Headquarters. From interviews, however, it is apparent that there has not been substantive communication about the relationship between CRING and COAR/SMQ reporting, either internally at Headquarters or outward with the field. There is some clear apprehension on the part of regional and country office staff that there will be duplication between CRING and the COARs, and a desire that the two efforts be reconciled in some fashion to prevent duplicative reporting requests. It was not feasible within the time allotted for the evaluability assessment to validate or refute this concern by analysing the degree of overlap between CRING and the SMQs.

It should be noted that many of the gaps mentioned in this report have already been recognized and are in the process of being addressed. Development and refinement of various systems include a revision and data cleaning exercise to improve tracking of resources linked to results; extensive revision of the country office annual reporting system to better align with the results of the Strategic Plan; and revision of the RAM 2.0 version to include a direct link to indicators. These systems, along with others, are contributing to filling gaps in the data needed to monitor, report on and evaluate the Strategic Plan. One important, outstanding issue is the manner in which the SMQs are phrased with some outcome areas asking specifically about results related to UNICEF's involvement and others making no such reference. Once all system improvements are fully operational, it will be necessary to revisit those systems to see how well they are being used to report on the priority results and to link resources with results. Particularly at the output level, there remains scope to objectively validate the indicators and monitoring systems to improve reliability.

Key finding 8: Data sources and systems for indicator measurement are identified and either in place or being established at the time of the evaluability assessment. Many lessons learned have been taken on board as these indicator measurement systems were rolled out and important steps taken to further refine and improve them. Due to timing, the evaluability assessment was not able to examine a full cycle of reporting that included these revisions. Some areas of concern remain with the SMQ responses serving as the basis for the majority of performance indicator reporting.

4.5 Systems to link resources to results

A substantial amount of effort has been devoted to allowing UNICEF to better align its resource allocation with results as elaborated in the Strategic Plan, notably in the integrated results and resources framework. Based on document review⁴⁰ and interviews, it appears that

³⁹ UNICEF Data and Analytics Section, CRING 2014 Guidance Note, CF/EXD/2014-001, 7 March 2014, Annex 1.

⁴⁰ United Nations Children's Fund, Programme Instruction: New Programme Information Database

the VISION system utilized by UNICEF will allow linkages between *intended* resource allocations and *anticipated* results with the following caveats:

- The system is based on the assignment of codes to specific results or activities and is therefore open to inter-coder interpretation and judgment.
- For country offices, codes are assigned to either country programme outputs or activities, which have an indicative or intended resource allocation. These intended allocations differ from budget codes and do not reflect actual expenditures.
- Items coded are anticipated results and not achieved results.
- Several of the attributes to be coded include categories that do not appear in the Strategic Plan. In those cases, the coding structure will not allow a one to one correspondence with components of the Strategic Plan.

Tables 15 and 16 demonstrate the resource allocation tracking system as it would be applied to a country office.

Table 15 shows that at the highest level, country offices are requested to examine and code their outputs along these four attributes. In sum, this table shows that country office-programmed resources can be aligned with strategic plan outcomes and programme areas as found in the Theory of Change document. In addition, country programme outputs are assigned a marker for gender equality and humanitarian action. The evaluability assessment did not examine how the four-point marker system was related to resource allocation, if at all.

Table 15: Attributes used to code country programme outputs

Attribute	Description	Relation to the Strategic Plan
Strategic Plan outcome code	Links the country programme output to one or more of the seven outcome-level results of the Strategic Plan	Directly related but country programme outputs can be coded to multiple strategic plan outcomes, thereby complicating 1:1 correspondence
Programme area code	Links each country programme output to one or more categories of programming activity under each outcome	These codes roughly correspond with key programme areas under each outcome in the Theory of Change
Gender Equality Marker	Each country programme output receives a single rating of expected contribution of the result to advance gender equality and/or the empowerment of girls and women using a four-point scale	Provides an assessment of gender content across Strategic Plan outcomes (categorical not numeric)
Humanitarian Action Marker	Each country programme output receives a single rating of expected contribution of the result to meet the CCCs using a four-point scale	Provides an assessment of humanitarian action across strategic plan outcomes (categorical not numeric)

Table 16 presents a lower level of resource tracking used in country offices. Whereas Table 15 shows the attributes assigned to country programme *outputs*, Table 16 presents the

(PIDB) Coding System in VISION for the 2014-2017 Strategic Plan, CF/PD/PRO/2014-001, 2014; United Nations Children’s Fund, ‘Results planning for 2014 and the new coding structure’, PowerPoint presentation; United Nations Children’s Fund, Attributes to Outputs at the time of Results based planning in VISION, Annex to CF/PD/PRO/2014-001, 2014.

attributes assigned at the level of country programme *activities*. While these additional codes provide some added clarity to resource flows, there are still substantial gaps between the attributes tracked and key features of the Strategic Plan. For example, the first attribute in Table 16 can link specific country programme activities with one of the seven strategic plan implementation strategies. However, the coding system also provides additional implementation strategies not included in the Strategic Plan, so the correspondence will not be complete. As another example, the third attribute, activity focus code, will identify country programme activities by determinant category (demand, supply, enabling environment and quality). This category has some similarities with Strategic Plan outputs 1-3, but again, cannot be assumed to accurately represent resource flows to those elements of the Strategic Plan.

Table 16: Attributes used to code country programme activities

Attribute	Description	Relation to the Strategic Plan
Generic implementation strategy code	Links each country programme activity to Strategic Plan implementation strategies along with additional categories	Categories go beyond the Strategic Plan so are not directly matched
Specific intervention code	Links each country programme activity to one or more types of programme activity categorized by Strategic Plan outcome and programme area	These detailed intervention codes correspond with key programme areas under each outcome in the Theory of Change
Activity focus code	Each country programme activity is coded according to three broad categories of: 1) demand, supply and enabling environment; 2) management/operations and 3) quality	Provides an assessment of the determinant category and an approximation of outputs 1-2; categories go beyond the Strategic Plan so are not directly matched
Humanitarian tag	Each country programme activity receives a 'yes/no' designation for whether the activity targets humanitarian action	Provides an binary (yes/no) assessment of humanitarian action across strategic plan outcomes

It seems that the systems of tracking resource flows do not yet entirely correspond to the major categories of the Strategic Plan at the level of outcomes or outputs. There is good correspondence at the level of programme areas and sub-areas but these are not directly included in the Strategic Plan results hierarchy. For the purposes of results-based management, it will be necessary to have (even estimated) resource flows by outcomes and outputs. It is understood that several outputs (i.e. outputs 4 and 5) are crosscutting in nature. Nonetheless, an agreed approach of apportioning those resources will be needed.

While UNICEF can categorize anticipated resources to elements of the Strategic Plan, it is not able to assess the adequacy of resources to achieve the intended results at this time. It will be difficult, if at all possible, to evaluate the ambitiousness of objectives and targets against available resources. Part of the difficulty is related to the integrated results and resources framework of the Strategic Plan, which is based on revenue and projected income, as opposed to a needs-based budgeting process that takes into consideration the results to be achieved. As with other aspects of the Strategic Plan, some sections are

attempting to address these gaps (see text box below).

Section efforts to link resources and the Strategic Plan results

The Child Protection Section undertook a global review of the funding situation for the 2014-2017 period. Although detailed information on the exact method was not available, there was a singular effort made to capture expected funding needs based on planned activities and funding actually available across Headquarters and regional and country offices. These data were then compiled to illustrate a complete global picture. Even at a high-level of approximation, the estimates from this exercise indicate that only about one third of required funds are available.

The Education Section has already begun a multi-year cost effectiveness analysis of the UNICEF education portfolio in order to inform future programming and planning. This analysis will delve into the available data and, ideally, will be used to improve the matching between indicators and results.

In part, the inability to link resources to results lies in the organization's means-based budgeting; that is, budgets are based on revenues supplied rather than an aggregate costing of work necessary to achieve results. In other words, the integrated results and resources framework is not based on amounts required to achieve results, but on the expected availability of resources. This places the organization at some risk as the results to be achieved are, in fact, not associated with evidenced-based resource requirements.

Key finding 9: Although current reporting systems categorize financial commitments by country office outputs, programme areas, key activities, sub-activities and implementation strategies, the materials reviewed did not specify how resource flows, across levels, would be tracked according to the results of the Strategic Plan (i.e. outcomes and outputs). Importantly, the adequacy of resources to achieve the anticipated results is not articulated.

5. Translating the Strategic Plan to regional and country office levels

One of the key aspects of the evaluability assessment was to identify the opportunities and challenges related to translating the Strategic Plan into country programmes of cooperation. This chapter addresses the process of rolling out the Strategic Plan, including the opportunities and challenges, and draws extensively from interviews conducted across levels (i.e. Headquarters and regional and country offices) and

It is important to bear in mind the underlying context that has been referred to throughout this document – namely that the Strategic Plan provides a results hierarchy that is intended to guide country programmes after consideration of national priorities, the situation analysis and the UNDAF. In addition, framing of outputs and outcomes and selection of specific areas of work derive from country dialogue and negotiation and not from a strict translation of the Strategic Plan.

This approach is evidenced in the guidance for the development of new country programmes that was updated soon after the Executive Board approved the Strategic Plan.⁴¹ Those guidelines stated that the “Strategic Plan 2014-2017 provides the global framework for country programmes” and that while countries should draw from the Strategic Plan, “the structure country programmes and corresponding outcomes will primarily derive from considerations of context, including the UNDAF and national development plans.”

Initial dissemination of the Strategic Plan from Headquarters to the regional level was facilitated, in part, by orientations (via presentations) and webinars. A set of introductory presentations⁴² developed by Headquarters for regional offices outlined the following three components to support country offices: 1) revised programming guidance and technical guidance, improve support processes; 2) identification of which posts/staff are responsible for leadership and quality of different elements of the Strategic Plan; and 3) establishment of an interactive knowledge exchange mechanism within the organization and with partners, using the structure of the Strategic Plan as an organizing framework. Of these, the evaluability assessment was able to identify that the first (i.e. revised programming guidance) has been created.

For the purposes of introduction, these materials appear to have been sufficient with a number of regional staff reporting that orientation materials were well received (e.g. one respondent stated: “I felt the ground was very well prepared to engage with the Strategic Plan”). Beyond the purpose of introduction however, these efforts seem to have limited scope (e.g. orientation

⁴¹ United Nations Children’s Fund, Programme Instruction: Updated guidelines on the development of country programmes to be approved by the 2014 Executive Board, CF/PD/PRO/2013-005, 8 November 2013.

⁴² UNICEF, ‘The UNICEF Strategic Plan 2014-2017: Realizing the rights of every child, especially the most disadvantaged’, PowerPoint presentation; The UNICEF Strategic Plan, 2014-2017. Operationalizing the New Strategic Plan; Jeffrey O’Malley, Looking forward: Post 2015, UNICEF 3.0/4.0 and the Strategic Plan.

to the Strategic Plan but not measurement of the Strategic Plan). Staff members in at least two regions described limited communication from Headquarters to the regional office specifically aimed at helping to guide country offices. As noted in several interviews, limitations in this regard included:

- From the level of the regional office to the country office, orientation typically reaches the deputy representative level (e.g. through Regional Management Team meetings) but doesn't necessarily trickle down from there; and
- Products didn't address the very limited knowledge at the country level of the Theory of Change.

In addition, several regional advisors received introductions and updates to the Strategic Plan through their respective Headquarters sections (e.g. WASH) or through the sector's regional network meetings. Technical staff at the regional office level commended the information conveyed through these means.

An additional element of rollout was the guidance issued for the development of regional office management plans.⁴³ These plans were intended to outline how units would fulfil their respective roles and functions to enable the collective achievement of the results in the Strategic Plan. Starting in early 2013, regional offices were advised to begin developing their plans based on a draft outline of the Strategic Plan. In September 2014, regional offices submitted revised regional office management plans covering the period 2014-2017. These plans (four of seven regions were reviewed) describe the planned regional contribution towards the global Strategic Plan within the context and priorities of that region. Examples of these stated priorities appear in Annex 9.

In several regions, region-specific priorities and agendas have been created. These are seen as a conduit for the translation of the Strategic Plan to country programmes. Although not necessarily anticipated when the Strategic Plan was rolled out, these regional agendas are viewed as a means for country offices to coalesce around a coherent programme with more focused priorities. Regional agendas typically result from discussion of priorities across the country offices in the region and alignment of those priorities within the framework of the Strategic Plan. Regional management teams (i.e. teams comprised of country representatives and the regional director) play an important function in the development, oversight and monitoring of these regional agendas.

Depending on the region, agendas might be supported through accountability frameworks, programme performance dashboards, theories of change, advocacy guidance, and/or multi-country evaluation plans. Some country office staff members that were interviewed appreciated the clarity and focus brought by the regional agenda, as illustrated by this deputy representative's opinion: "The regional office has taken what was done at global level and made it more relevant to us; they broke it down for us and without that we would have been all over the place."

Across regions, regional office staff provided a significant amount of orientation to country representatives and deputy representatives. As described below, in addition to orientation, the regional offices provided substantial support to country offices to align with the Strategic Plan during planning and review exercises (e.g. strategic moment of reflection, mid-term reviews and country programme development). As noted by more than one regional office

⁴³ United Nations Children's Fund, Programme Instruction: Guidelines for the Preparation of the 2014-2017 Office Management Plans, UNICEF, CF/PD/PRO/2012-003, 2012.

planning, monitoring and evaluation staff member: “We see it as our job to help the country offices adapt the Strategic Plan to their context.”

The country programming cycle provides several opportunities for country offices to consider the Strategic Plan within the context of national priorities. As outlined below, these opportunities include the development of the country programme, mid-term reviews, strategic moments of reflection and annual reviews.

The beginning of a country programme cycle, which is usually five years in duration, provides the most visible opportunity for translating the Strategic Plan. The CPD and the CPAP guide the programme cycle. Current guidance for the development of these documents provides a light touch reference to the Strategic Plan and how it is to be considered in preparation. In preparing a CPD, country offices are to describe their programme focus in terms of national development priorities and UNICEF’s contribution to the UNDAF and are advised that: “description of programme components and outcomes should be guided by the Strategic Plan 2014-2017.”⁴⁴ The only documented verification that this step occurs was found in a recently developed ‘Checklist for the Review of Draft Country Programme Documents,’ published in October 2014. This checklist includes a single criterion on alignment with the Strategic Plan, as follows: “Alignment with the Strategic Plan 2014-2017 is evident.”

In terms of tangible linkages between country programmes and the Strategic Plan, country offices are instructed to use the Strategic Plan’s seven implementation strategies when articulating the rationale for and mix of programme strategies to be employed. In preparing the results and resources framework, country offices are required to identify (by number) which of the Strategic Plan’s outcome(s) align with the country programme outcomes. It is notable that no mention is made of monitoring and evaluation requirements vis-à-vis the Strategic Plan nor is the summary budget format structured by the Strategic Plan’s outcomes.⁴⁵ In other words, country offices are not advised to adopt indicators from the Strategic Plan’s results framework nor to align country office budgets to strategic plan result categories.

MTRs are another opportunity to translate the Strategic Plan into country programmes, although an MTR would generally result in a more incremental adjustment than a CPD. Recent guidance has made the MTR an optional exercise. Nonetheless, it is still utilized as a tool for reviewing and adjusting country programmes, though the form may differ from what it was previously. For example, one regional office planning, monitoring and evaluation staff member explained that country offices have a choice of doing a full MTR, an MTR light, no MTR or a strategic moment of reflection.

An example of the use of the strategic moment of reflection comes from UNICEF Zimbabwe. With the current CPD coming to an end in 2015, UNICEF Zimbabwe developed strategy papers in line with the Strategic Plan’s seven outcomes and developed a first draft of the CPD. A strategic moment of reflection was used in a forward-looking manner to review and validate the proposed strategies for the Country Programme 2016-2020. An example of the use of the

⁴⁴ United Nations Children’s Fund, Programme Instruction: Revised Format for Country Programme Documents submitted and approved at the second regular session of the Executive Board 2014, UNICEF, CF/PD/PRO/2014-002, Addendum to CF/PD/PRO/2013-005, 2014.

⁴⁵ Ibid. Note: Guidance for the preparation of country programmes in 2014 used the following main categories for summary budgets: health and nutrition; WASH; basic education; child protection; social inclusion; and cross-sectoral.

MTR comes from UNICEF Bangladesh.⁴⁶ Drawing on the analysis conducted within the framework of the MTR, UNICEF Bangladesh chose to revise the structure of the Country Programme 2012-2016 to further align with the Strategic Plan and UNDAF. Changes will be reflected primarily through the revised integrated results and resources framework 2015-2016.

Almost every regional office reported either continued use of MTRs (albeit modified) or strategic moments of reflection to consider the Strategic Plan and make adjustments.⁴⁷ There is considerable variation between regions in terms of how opportunities are being pursued. For example, where one region is making extensive use of strategic moments of reflection in advance of developing a new CPD, another region might view a strategic moment of reflection as having no real traction.

Other avenues for translation included country office annual reviews and input into national planning processes. The latter was cited particularly in cases where middle-income countries have the required human and financial resources and fairly well developed national strategic plans with specific gaps. In this regard, it was found helpful to have the Strategic Plan structured by sectors, which is seen as well aligned with the structure of national programmes and priorities. In regards to the structure of the Strategic Plan, one regional office planning, monitoring and evaluation staff member said “It makes it easier for us to articulate the structure of our Strategic Plan with that of the national response. Even in countries that were planning in late 2011, components can be aligned according to the sector. The real challenge is to align the results in national programmes.”

Finally, additional strategy supportive pieces (e.g. research priorities and policy advocacy priorities) are under development at Headquarters and may represent important milestones for UNICEF in terms of alignment with the Strategic Plan.

Of course, alongside these opportunities are challenges related to translating the Strategic Plan to country programmes. This chapter outlines some of the types of challenges that emerged most frequently, namely communication and the timing of the rollout of elements of strategy implementation.

In terms of communication, most regional office staff reported that they had multiple opportunities to provide input to the development of the Strategic Plan. However, these opportunities were often characterized as rushed with little if any feedback or response to their inputs. A few country office staff reported that they were actively engaged via their regional office counterparts. Some staff members interviewed did not themselves have the opportunity to directly engage in the strategic planning process but were confident that others in the regional office did have that chance.

In contrast, no regional office staff members reported that they had provided input into the Theory of Change document and several had not seen the Theory of Change document at the time of the evaluability assessment interviews (late October through early December 2014). One regional office planning officer described ample opportunity to provide input to the Strategic Plan and contrasted that with the results framework and Theory of Change

⁴⁶ United Nations Children’s Fund, ‘Mid-term review of Government of Bangladesh-UNICEF Country Programme 2012-2016’, Ruposhi Bangla Hotel, 29 June 2014.

⁴⁷ Some regional office staff regret the loss of the MTR as a tool that prompted a larger process of research and review and a platform for sharing findings. For example, the MTR provided an occasion for rigorous barriers and bottlenecks analysis to inform new country programme development.

documents, which they viewed as developed at Headquarters without much regional input. This point of view was consistently heard from the regional offices.

The time involved in developing and rolling out the Theory of Change and results framework⁴⁸ appears to have been substantially underestimated, with some lengthy delays as a result. One important factor is the iterative process through which the Theory of Change was developed, with a first draft available in June 2013 (before the approval of the Strategic Plan) and the final version available nearly a year later in May 2014.

Similarly, guidelines for conducting annual reviews/reporting were to be developed by 31 October 2013 and additional technical guidance on indicators, including setting baselines and targets for the new Strategic Plan, was expected by 30 November 2013. The annual reporting system intended to create baselines for the Strategic Plan was launched in early 2014. As noted above, the information that returned was determined to be of low quality and ultimately unusable, primarily due to lack of consistent definitions and standardization. The methods for collecting indicator data (on output-level measures generated from country offices) were significantly revised, more standardization was provided in the form of detailed indicator sheets, and the exercise was re-launched in November 2014.

However, at the time of the evaluability assessment interviews, most regional planning, monitoring and evaluation advisors had not seen the revised results framework indicators and were unaware of the detailed indicator sheets. Granted, these materials were distributed in November 2014 so some interviews may have preceded distribution. Some concern was expressed as one year of the strategy period had passed with no published guidance on revised indicators. At a fundamental level, there was lack of clarity as to whether or not country offices were supposed to include results framework output-level indicators within their own country programme resources and results framework. A respondent from one country office explained that: "It's not possible to be completely in alignment with the Strategic Plan and the national reality. We are still in negotiation with the regional office but I think we will end up with 20 to 25 per cent of the indicators for our CPAP in line with the results framework. Some of the indicators are not possible for us to collect."

Work is currently underway to develop and issue more technical resource materials to country offices as an additional guide to programming. Structured around each outcome area, these materials will provide an overview; short descriptions of each output and key programme areas; and explanation of the Theory of Change for the outcome, including available evidence and how that evidence base fits within the Theory of Change. These materials will be accompanied by video and audio prepared by the Division of Communication. The resource materials are expected to be complete by the end of February 2015. Although this type of resource package will be an important supplement to the Strategic Plan, if available as intended, almost one third of the period covered by the Strategic Plan 2014-2017 will have passed.

In addition, counter to initial impressions, opportunities are not limited to the development of country programmes but also include mid-term reviews, strategic moments of reflection and annual reviews. These opportunities have not been leveraged as uniformly as they might have been due to uneven communication around the Theory of Change and delays in rolling out elements of Strategic Plan-related results reporting. Other rollout needs relate to regional

⁴⁸Other supportive dimensions such as ICON and the public website are still largely structured according to the previous MTSP.

offices' views of limited country office capacity to produce good theories of change and the need for guidance from Headquarters to help country offices strengthen their theory of change thinking.

Key finding 10: Country offices, with the strong support of regional offices, have already utilized a number of opportunities in the country programme cycle to reflect on the Strategic Plan and make adjustments in their programmes. Although not outlined in the Strategic Plan, regional offices have facilitated the uptake of strategic plan elements in country programmes by interpreting the global Strategic Plan into region-specific priorities.

6. Suggestions for future evaluations

The fourth objective of the evaluability assessment is to provide guidance on approaches to evaluations of the Strategic Plan with a view to enhancing its utility. Three questions were identified under this objective, as follows. Is there an evaluation, research, learning agenda, strategy or plan in place for each outcome? Similarly, is there such a strategy or plan in place for each crosscutting theme and normative principle? If so, do these fit into any overarching learning strategy for the Strategic Plan? Finally, how might the utility of planned evaluations be enhanced? Answers to these questions were sought through interviews and document review.

The evaluability assessment found that evaluation and research plans framed around the Strategic Plan provide a foundation for such an agenda. As a first element, a plan for global thematic evaluation 2014-2017⁴⁹ was designed specifically to evaluate the main building blocks of the Strategic Plan. As such, the plan for global thematic evaluations includes evaluations devoted to all seven outcomes and crosscutting themes (i.e. gender equality and humanitarian action). Implementation strategies are also subject to evaluation and incorporated into the plan for global thematic evaluations with the exception of those thought to be better examined through regional- or country-level evaluations (i.e. service delivery and south-south cooperation). The plan for global thematic evaluations aims to provide evaluation evidence on the relevance, efficiency and effectiveness of UNICEF's work under the Strategic Plan.

In addition to the global thematic evaluation plan, the organization has created a set of research priorities structured around the Strategic Plan and its outcomes.⁵⁰ Importantly, this is the first time that UNICEF has prepared organization-wide research priorities.⁵¹ To a large extent, the identified priorities are pegged to evidence gaps in the Strategic Plan, by outcome area, and further examine key aspects of equity. In the case of the health outcome area, research will also be formulated to support data generation and evidence regarding the impact, outcome and output indicators as outlined in the Strategic Plan. A large number of crosscutting research priorities have also been identified, both within outcome areas and in a separate category.

⁴⁹ United Nations Children's Fund, Plan for global thematic evaluations, UNICEF Executive Board first annual session 2014, Item 8 (b) of the provisional agenda, 13 November 2014, <http://www.unicef.org/evaluation/files/2014-3_Global_evaluation_plan_2014-2017_DL-HR.pdf> accessed 12 August 2015.

⁵⁰ United Nations Children's Fund, *A Framework for Research, 2014-2017*.

⁵¹ In cross-reference to a point made in section 3.3 of this report, some programme areas that figure prominently in the Theory of Change and are formulated as global-level performance indicators are themselves the subject of priority research to fill gaps in the evidence base. These include: Output 3/WASH (sustainability compacts), for which priority research questions include: Has the sustainability of water services improved as a result of the implementation of the sustainability compacts and related sustainability checks? and Output 1/WASH (water safety plans), which is to be examined in terms of: What evidence exists to demonstrate improved protection of public health from the implementation of water safety plans? To what extent does the introduction of a water safety plan lead to improved management capacity and practices?

Currently, there is no overarching learning strategy for the Strategic Plan. However, taken together, the global thematic evaluation plan and the research priorities framework provide good coverage of the main results identified in the Strategic Plan. In particular, there is good coverage of outcomes and crosscutting priorities related to gender equality and humanitarian action and equity. In terms of normative principles, gender equality is covered (as a crosscutting theme) and several research questions have been identified for environmental sustainability. Research related to rights appears throughout the set of priority topics. A challenge to the agency will be to manage and integrate these activities, be they independent evaluations or programme-focused research, into a comprehensive evidence base for the Strategic Plan. If this knowledge base is to lead to action, then at a minimum, it should be packaged and made available to key audiences for application. It is particularly important that evaluation and research on the crosscutting areas are viewed and managed in a comprehensive and integrated manner.

At the time of the MTR, progress towards the global thematic evaluation plan and research framework should be carefully assessed. At that point, a determination should be made regarding the nature of any final evaluation of Strategic Plan implementation. At least two options are possible. The first option would build on a robust and timely set of evaluation and research activities as the prime evidence base. Those priority issues and questions that have not been adequately addressed by that point would represent the evidence gap. In this option, a final evaluation would take the form of a meta-review and synthesis drawing from the available information, generating new findings to fill the gaps and providing an overarching summative analysis. The Global Fund to Fight AIDS, Tuberculosis and Malaria is currently employing this approach for their Strategic Review 2015. A second option would be triggered if, at the time of the MTR, it is determined that the coverage of priority issues and questions through existing evaluation and research is not adequate or has not kept pace with strategy implementation. In this scenario, a more standard final evaluation may be needed as the magnitude of the gaps in the evidence base will be larger.

The final question related to the how the utility of planned evaluations might be enhanced is considered below. Regardless of the use of options outlined above or others, these recommendations may help steer those efforts. These derive both from the key findings as well as other observations made during the evaluability assessment.

- In simple terms, a significant challenge to the evaluability of the Strategic Plan in principle is that it looks like a plan but is applied like a framework. Accordingly, the Strategic Plan is most productively evaluated as a framework, with a significant role for regional-level mediation to help guide country-led programme priorities. Given UNICEF's decentralized decision-making, it would be a mistake to evaluate the Strategic Plan as one that is firmly applied across levels or strongly directional.
- The actual translation of the Strategic Plan to country programmes will not be evidenced in any form of adherence to specific outcome or output statements or wholesale adoption of programming approaches. A challenge to evaluation will be to identify how and to what extent the Strategic Plan has guided dialogue and country level decisions in a nuanced manner. Evaluation of the Strategic Plan should conceptualize and examine how a global framework adds value to country programmes in widely varied contexts.
- Likewise, assessment of the causal chain will be best informed by systematically looking at and documenting strategy implementation across a range of countries (e.g.

fragile, middle-income, high burden) where a rich and diverse set of results can be identified but not easily aggregated. The MTR should examine implementation in a broad range of countries. The MTR and final evaluation of the Strategic Plan translation to country programmes must also take into account how regional priorities and agendas have acted as mediating factors.

- In regards to the results hierarchy, there are several cases where the descriptions at the output level are concrete enough that they represent a testable element of the Theory of Change.⁵² Even in these cases, there is no causal chain described but there is sufficient clarity on what types of actions UNICEF will take to achieve the output. While evaluation questions could be framed to plausibly link or test the association between these activities and the output, attention is likely better focused on the relationship between each section's technical theory of change and the corporate Theory of Change.
- The global Situation Assessment and Analysis and the MTSP end-of-cycle-review provide a critical knowledge base for evaluating the relevance of the Strategic Plan. These materials outlined new areas of strategic importance and emphasized the assessment of new programme areas based on comparative advantage, complementarity and available resources. While such strategic choices were clearly made, no consolidated record was available that described how these strategic choices were weighed and taken. Such a record would likely need to be created for a complete evaluation of relevance.
- In terms of the forward-looking goals, there is a high degree of correspondence between the SDGs and the Strategic Plan's core components. The global commitments that will be made regarding the SDGs also provide an opportunity for forward-looking evaluation. Looking at standards related to the evaluation of normative work,⁵³ UNICEF might consider an evaluative framework that prospectively examines the agency's role in supporting governments and others to integrate the SDG-related norms and standards into legislation, policies and development plans and their implementation.
- The Strategic Plan not only positions humanitarian action as a crosscutting theme, but also clearly states the intent to capitalize on UNICEF's agency in the areas of development and humanitarian action through greater integration. This element of the Strategic Plan is represented in the global thematic evaluation plan (i.e. as an evaluation focused on resilience), the research framework (i.e. researching the alignment and linking of humanitarian and development programming) and some work is already underway wherein the WASH section, together with the World Bank, is launching a study to identify the gaps (silos) between development and humanitarian WASH and recommend actions to reduce the gaps. The commitment to integrate humanitarian action and development programming should be the subject of evaluation as well. The planned evaluation on resilience should not be narrowly focused but should encompass humanitarian action as a crosscutting theme. In addition to questions about effectiveness, there are important process-level issues to

⁵² Examples include Output 2/HIV/AIDS, Output 2/nutrition and Output 3/WASH.

⁵³ United Nations Evaluation Group, 'UNEG Handbook for Conducting Evaluations of Normative Work in the UN System', November 2013.

be examined regarding the organizational changes being brought about.⁵⁴ It would benefit the organization to have a rigorous and independent examination of how UNICEF intends to bring about this integration and the progress that is being made.

- Also arising from the evaluability assessment is an element that cuts across outcomes and may be considered for inclusion in select future evaluations. This element deals with the ability of each outcome area to estimate the actual resources needed to achieve certain results. Currently, the Strategic Plan is budgeted based on anticipated revenues in line with assumptions about past levels of income. This places the organization at some risk, as the results to be achieved are not in fact associated with evidence-based resource requirements. One way of addressing this gap is by embedding an objective into planned evaluations to assess the ability of the programme being evaluated to cost results. In addition, several interviewees with strong analytical skills pointed to the wealth of costing data available within UNICEF and the meagre use of this data. For example, there is at least a suggestion that some of the basic components needed for a value for money assessment may be on hand and underexploited.
- To be fully evaluated, rights-based approaches will need to be conceptualized as integral to all outputs and outcomes (e.g. support for rights holders in Output 1). There may be an underlying dichotomy to be resolved regarding the treatment of rights as a programming approach (as per the human rights-based approach to programming) versus an output-level result related to increased capacity of duty bearers. Likewise, efforts to evaluate the Strategic Plan will necessitate a crosscutting examination of gender-focused programming. Limiting the evaluation of gender-focused actions to this single output would be a severe oversight on the part of evaluators. For the purposes of evaluation, the Gender Action Plan should be viewed as a critical accompaniment to the Strategic Plan. A planned evaluation of the Gender Action Plan should not overlook the critical linkages between that plan and the crosscutting gender component of the Strategic Plan.

⁵⁴ The following suggestions emerged from interviews regarding themes to be evaluated as part of the organizational changes being put into place: To what extent have divisions (e.g. Programme Division, Division of Data, Research and Policy, Public Partnerships Division and the Division of Communication) included humanitarian action in their own management plans and work plans? What results have emerged from efforts to strengthen humanitarian action as part of the efficiency and effectiveness agenda? What are regional- and country-level changes? How have regional directors and deputy regional directors been involved in strengthening integration? Have human resources for humanitarian action been strengthened? How much has been invested in this? Can UNICEF report on how regular resources and other resources have been used for humanitarian action?

7. Implications for UNICEF

This chapter presents implications that emerged from the evaluability assessment by examining key findings in terms of what these findings mean for the organization as a whole and advancing a set of recommendations. These recommendations take into account the actions needed in the near future to make the current plan more evaluable in view of the MTR planned for late 2015 and are also geared towards helping the organization as it prepares the next strategic plan. Chapter 6 addressed forward-looking suggestions on how to evaluate the Strategic Plan.

The analytical lens used to assess the current Strategic Plan examined elements in place taking into account relevance, plausibility, structural and contextual issues. A strategic plan is a product of a planning process that defines an organization's medium- to long-term goals, as well as the resources required to achieve them. A strategic plan serves to establish and implement a vision within an organization and to support the organization to make choices and respond to emerging issues, needs and mandates. A strategic plan should contain the following elements: a) clear objectives; b) direction and focus; c) well-defined principles or normative issues and standards for measuring them; d) articulation of the implementation strategies; e) alignment of capacity and resources to match with objectives, focus and implementation strategies; and f) explicit identification of results and indicators for measuring progress.

The assessment finds that the objectives and direction of the Strategic Plan are clear. It notes that the standards for measuring normative issues are weak. The assessment highlights the lack of specifically stated linkages between outputs and outcomes and between outcomes and impact. This lack of clarity about how each outcome links to the organization's intended impact means that the Strategic Plan cannot be used to make trade-offs or weigh decisions across the outcome areas should the question or necessity arise. The relationship between outputs and outcomes is equally problematic as the lack of specifically identified linkages means it is not possible to establish how outputs contribute to outcomes.

The assessment notes that UNICEF cannot currently determine the adequacy of resources to achieve the anticipated results. This challenge lies not in the reporting systems but follows from the use of a budgeting approach that is means-based (based on revenues supplied) rather than results-based (providing an aggregate costing of work necessary to achieve results). This places the organization at some risk, as the results to be achieved are not associated with evidence-based resource requirements.

The evaluability assessment finds that the Strategic Plan serves as an overarching framework. The analysis notes that while the Strategic Plan looks like a corporate-level plan that articulates a results structure from outputs to impact, it is applied as a framework and widely considered as such by UNICEF staff. The reason for this lies in the UNICEF modus operandi, which is to support national governments. Programme planning occurs in negotiation with governments around the world. Therefore, UNICEF does not have complete control over its programming at the country-level. The results UNICEF seeks are usually achieved in partnership with governments, as well as with non-governmental organizations and civil society.

The key question is how to determine if UNICEF is performing well. Therefore, if the Strategic Plan transitions to a framework, mechanisms must be found to aggregate results meaningfully at both the country and global levels. Results must remain central to the plan or framework.

Recommendations

The following recommendations are based on the findings and conclusions presented within this report:

1. Reorient the Strategic Plan in line with its application as a framework.

The evaluability assessment found that the Strategic Plan is articulated logically as a plan but is applied in practice as a framework to better address national development priorities and contexts. Recognizing this, it is recommended that the next round of strategy development should more explicitly account for the need for greater flexibility and adaptation (e.g. more flexibility at the level of outputs, reduced use of aggregated results reporting and greater recognition of context-specific variables). This should be fully considered and negotiated with key stakeholders, taking into consideration the UNICEF modus operandi as a highly decentralized organization. The upcoming MTR should, in addition to assessing strategy implementation progress, further examine the implications of using a strategic framework approach for the 2018-2021 period and, accordingly, provide an indication of next steps.

2. Newly drafted CPDs and CPAPs should be reviewed in order to better understand how they relate and contribute to the Strategic Plan.

With substantial regional office support, country offices are making efforts to adjust country programmes to be better aligned with the Strategic Plan. Greater alignment of priorities and clarity of results could be facilitated through the review of newly developed CPDs and CPAPs and sharing of lessons. This review process should aim to understand how well country offices have used determinants analyses and other tools to focus programmes on the rights of the most disadvantaged and excluded children in the country context; and in turn, how well these country-determined foci align with results articulated in the Strategic Plan.

3. Clarify results, measurement approaches and indicators for equity.

Equity is a central tenet of the UNICEF mission and thoroughly integrated into the Strategic Plan. However, there is a lack of clarity in the Strategic Plan on defining results relating to equity and how these will be measured. It is not necessary for UNICEF to have just one single definition of equity. However, all equity concepts that are employed need to be operationalized with sufficient clarity and rigor to allow measurement and evaluation.

4. Conduct a review and adjust, as needed, indicators and systems to measure progress towards the results of the Strategic Plan.

Extensive investments have been made in systems to report on the Strategic Plan indicators. At higher levels of the results chain (i.e. impact and outcome), UNICEF is a central contributor to global data sets, analytical work and methods development. Concerns

related to indicator quality at these levels are limited. Performance indicators at the level of outputs are of more variable quality and, despite improvements, there are questions about the systems used to generate data. It is recommended that an assessment of the quality of the output-level performance indicators be conducted with particular attention to those generated through the country office annual reporting system. Any further adjustments to indicators should be based on objective assessment of the performance of this system. The review should be explicit about how results should be measured both at country and global levels.

5. Conduct an exercise to determine the need and feasibility of moving the organization towards budgeting practices that are more supportive of results-based management that enable better linking of resources and results.

Currently, the Strategic Plan's integrated budget relies on projections of resource availability (i.e. means-based budget). The organization should explore various ways to link capacity and resources with results.

