Global Evaluation Reports Oversight System (GEROS)

JANUARY 2013

unite for
children

unicef 🙂

Introduction and justification

The evaluation function seeks to strategically contribute to UNICEF's performance by providing good-quality evidence for learning, decision making, policy advocacy, as well as accountability purposes. Reflecting the decentralized nature of UNICEF, the majority of evaluations supported by UNICEF are managed at a decentralized level. While the decentralized nature of the evaluation function ensures that evidence generated is relevant to the local context and therefore more likely to inform national policies for children, it poses the challenge of setting up a consistent corporate system to ensure good quality and credibility.

To face this challenge, the Global Evaluation Compact committing Evaluation Office and Regional Offices to work jointly to strengthen the evaluation function was endorsed in 2009. The Compact focuses on: enhancing strategic planning of evaluations through prioritization of major evaluations; promoting and supporting quality of evaluations; improving use of evaluations and management response; strengthening internal evaluation capacity; developing national evaluation capacity in support of country-led M&E systems.

As a result, Key Performance Indicators (KPI) improved:

- Submission of evaluation reports to the Global Evaluation Database reached 98% in 2012 from the 2008 baseline of 20%.
- Submission of Management Responses to the Global Tracking System reached 93% in 2012 from the 2009 baseline of 10%.
- Implementation of Management Responses increased from 62% in 2011 to 82% in 2012

However, while significant progress has been made in the above mentioned KPI, in 2012 the percentage of good quality evaluation remains stable at 41%, those almost satisfactory increased to 34%, and the percentage of bad quality evaluation decreased to 24%.

In 2009, the Global Evaluation Reports Oversight System was designed in the context of the Global Evaluation Compact. After three years of implementation, in 2012 GEROS went through a rapid review involving COs, ROs, EO and the external company that won the bidding for the period 2012-2015.

Purpose

The Global Evaluation Reports Oversight System (GEROS) has four main objectives:

- <u>provide senior managers</u> with a clear and short independent assessment of the quality and usefulness of individual evaluation reports, including those commissioned by their own offices;
- <u>strengthen internal evaluation capacity</u> by providing to commissioning offices feedback with practical recommendations on how to improve future evaluations. Commissioning offices can also use the feedback received to better assess the performance of external consultants to be hired for future evaluations.
- <u>report on the quality of evaluations reports</u>, by reviewing and assessing the quality of final evaluation reports commissioned by UNICEF Offices. Quality of evaluations reports is reported to senior management mainly through three channels: a) annual report of the EO Director to the Executive Board; b) the Global Evaluation Dashboard, and c) inclusion of this information in the Global Evaluation database;
- <u>contribute to corporate knowledge management and organizational learning</u>, by identifying evaluation reports of satisfactory quality to be used in meta-analysis to be shared within the organization, as well as facilitating internal and external sharing of satisfactory evaluations reports

GEROS and UNICEF evaluation function

GEROS is an <u>organization-wide system.</u> The assessment of final evaluation reports, managed by the EO Systemic Strengthening unit, is complemented by quality-assurance mechanisms designed and implemented by COs and ROs, which focus on reviewing draft ToR and draft reports against the UNICEF/UNEG ToR and Report checklists, giving real time feedback to allow quality improvement of final ToR and reports.

GEROS uses the <u>UNEG evaluation reports standards</u> as basis for review and assessment of final evaluation reports, while ensuring specific standards relevant to UNICEF are taken into consideration (Annex 3).

To ensure credibility and effective reviews, the assessment of final evaluation reports is <u>outsourced to an external and independent high-quality experienced company</u> selected through a transparent and open bidding process.

While the Global Evaluation Database contains evaluations, surveys, studies and researches, GEROS reviews 100% <u>evaluation reports uploaded by COs to the Global Evaluation database by 31st march of the following year</u>. EO Systemic strengthening unit screens reports uploaded to the Global Evaluation Database using PPP Manual definitions (Annex 1), and submits the reports identified as "evaluations" to the external company that makes a final screening and reviews only evaluation reports. EO uploads reports in the UNICEF Internet webpage; carries out meta-analysis of selected priority areas, and present them at the Executive Board; and informs appropriate HQ Divisions/ROs of the good quality reports in their own area of work/regions. Each Division/RO is expected to make any further meta-analysis as necessary, including extrapolating good practices and lessons learned.

What constitutes a highly satisfactory evaluation report

An evaluation report is assessed as highly satisfactory when it is a <u>credible report that addresses the evaluation purpose and objectives based on evidence, and therefore can be used with confidence</u>. That is to say, a highly satisfactory evaluation report will provide a clear and complete assessment of the object of the evaluation, based on evidence compiled and analyzed in accordance with UNICEF-adapted UNEG standards, such that its conclusions and recommendations can be deemed to be credible and are thus a sound basis for decision-making.

Evaluation reports are reviewed using the UNICEF-adapted UNEG Evaluation report standards Matrix (annex 3) to assess the following <u>core elements</u>:

Well structured, logical and clear report

The report is logically structured with clarity and coherence (e.g. background and objectives are presented before findings, and findings are presented before conclusions and recommendations). It reads well and is focused.

Clear and full description of the 'object' of the evaluation

The report describes the object of the evaluation including the results chain, meaning the 'theory of change' that underlies the programme being evaluated. This theory of change includes what the programme was meant to achieve and the pathway (chain of results) through which it was expected to achieve this.

The context of key social, political, economic, demographic, and institutional factors that have a direct bearing on the object is described. For example, the partner government's strategies and priorities, international, regional or country development goals, strategies and frameworks, the concerned agency's corporate goals and priorities, as appropriate.

The evaluation's purpose, objectives and scope are fully explained

The purpose of the evaluation is clearly defined, including why the evaluation was needed at that point in time, who needed the information, what information is needed, how the information will be used. The report provides a clear explanation of the evaluation objectives and scope including main evaluation questions and describes and justifies what the evaluation did and did not cover. The report describes and provides an explanation of the chosen evaluation criteria, and/or other criteria used by the evaluators.

Appropriate and sound methodology

The report presents transparent description of the methodology applied to the evaluation that clearly explains how the evaluation was specifically designed to address the evaluation criteria, yield answers to the evaluation questions and achieve evaluation purposes.

The report presents a sufficiently detailed description of methodology in which methodological choices are made explicit and justified and in which limitations of methodology applied are included. The report gives the elements to assess the appropriateness of the methodology. Methods as such are not 'good' or 'bad', they are only so in relation to what one tries to get to know as part of an evaluation. Thus this standard assesses the suitability of the methods selected for the specifics of the evaluation concerned, assessing if the methodology is suitable to the subject matter and the information collected are sufficient to meet the evaluation objectives.

Findings, conclusions, recommendations and lessons learned are based on evidence and sound analysis

Findings respond directly to the evaluation criteria and questions detailed in the scope and objectives section of the report. They are based on evidence derived from data collection and analysis methods described in the methodology section of the report.

Conclusions present reasonable judgments based on findings and substantiated by evidence, providing insights pertinent to the object and purpose of the evaluation.

Recommendations are relevant to the object and purpose of the evaluation, are supported by evidence and conclusions, and were developed with involvement of relevant stakeholders. Recommendations clearly identify the target group for each recommendation, are clearly stated with priorities for action, are actionable and reflect an understanding of the commissioning organization and potential constraints to follow up.

Lessons learned are grounded in the evidence arising from the evaluation, but provide insights that are relevant beyond the specific scope of the projects, programs or policies evaluated. Lessons learned highlight elements of preparation, planning, design or implementation that can be expected to have positive or negative effects on performance, outcome, or impact.

The GEROS Methodology

The Basis of the Methodology

The Methodology is based on:

- The production of high quality, informative, consistent and concise reporting.

- Understanding that there are different levels, scales and scope of evaluations

- Giving special focus on evaluations of strategic significance to UNICEF as an organisation

- Adapting for joint-evaluations, including UNDAF evaluations, where UNICEF has collaborated with other institutions to deliver, as well as country-led evaluations.

The Review Process

Reports are tracked throughout the review process. While all reports are systematically and rigorously reviewed according to UNICEF-Adapted UNEG standards, 20% of them (i.e. those flagged by UNICEF as of particularly strategic significance to the organization, as well as those rated moderately satisfactory), benefit from a senior review as an additional layer of assurance.

Individual review process consists of three main parts:
- Part I includes the recording of key report details (such as title, reviewer, etc) that allow the report to be tracked. Additional data is collected at this point in order to allow future analysis of trends in the evaluation function. This data includes codes for:

    o Country / Region / Main Sector of Focus

    o Evaluation Scale / Size / Subject / Stage (a detailed taxonomy is presented later)

- Part II provides qualitative and scored feedback on the Six UNICEF-Adapted UNEG Standards , which are dealt with in separate sections in the review of evaluation reports. Each of these Standards and their corresponding sections are considered independently:

    o The reviewer answers a series of questions that have been derived from the UNICEF-Adapted UNEG Evaluation Reports Standards. These are designed to inform a qualitative story on the level of each of the six sections to be reviewed, noting any points that will subsequently inform the reviewers reflection on areas for future improvement in evaluation practice (to be captured as part of the 'Overall Rating' step for each report).

    o For each of the six sections, few *key questions* have been identified (marked by shading). These key questions are recommended as a starting point for the reviewer to begin their overall analysis of the confidence they have it that section.

    o Informed by process so far the reviewer provides overall qualitative feedback on the Section and applies a test to answer the *Cornerstone Question* for that Section according to a 4 point scale. This test is elaborated later.

    o Based on the scoring for the section and the analysis provided, the reviewer provides feed-back for UNICEF staff on how to improve future evaluations with respect to the section concerned including ways to address weaknesses and to maintain good practice identified.

    o The reviewer provides focused feed-back on the assessment of the section for Senior Management, including positives as well as negatives and shortly justifying the rating.

- Part III prompts the reviewer to provide qualitative feedback with questions related to the coherence and credibility of the entire report. Based on this process and the understanding gained from Part II, the reviewer is then required to apply a test to provide an overall rating for the report making use of a 4 point scale.

Quality Assurance

There are inherent challenges in managing the consistent application of qualitative criteria to a broad range of evaluations, across a range of reviewers. The following quality assurance mechanisms to ensure consistent interpretation of evaluation standards are applied:

- Selection of reviewers with deep knowledge of evaluation and understanding of the evaluation function in UNICEF.

- An initial training with all reviewers to read through and discuss all the standards, working with examples. This is led by a reviewer with the strongest knowledge of applying the standards.

- All reviewers assess the same example of a report. Ratings and feedback are compared among reviewers, differences discussed, and a definitive interpretation reached. This process can be repeated if necessary.

- Ongoing peer review of a sample of completed work from each reviewer. These peer reviews are to be conducted near to the outset of the review process in order to allow for early adaptation of ratings if differences are found. The review ratio is 20% of the number of reports reviewed, with a minimum of 2 peer reviewers per review. The peer-reviewed rating sheets are returned to the original reviewer along with the opportunity to revise their interpretation.

- Providing all reviewers with a good range and number of reports in their review sample: the experience of seeing many examples of evaluations assists in building a better awareness of how the range of ratings available equates to real life examples.

Once developed, this quality-assured capability can be maintained more easily and affordably: using the peer review of prioritised evaluations and those rated as "moderately satisfactory" (mentioned previously) as a check. However, if additional reviewers are brought into a review group, then it is important that they pass through a full quality assurance process as part of their orientation.

Four Point Scale

This methodology makes use of a four-point scale in order to force reviewers to assess reports as either good or not. The use of four-points, rather than two, also enables the identification of key clusters of reports:

- Outstanding reports that represent good practice (Outstanding Good Practice)

- Reports that are highly satisfactory

- Reports that mostly satisfactory (could be good quality with a little more work)

- Reports that are unsatisfactory

Horizontal comparability

This proposed methodology aims to enhance comparability across evaluation reports in multiple ways.

- By collecting coded evaluation typology and background data the analysis of batches of reports can be broken down in multiple ways according to what is required. Reports can thus be compared like-with-like rather than against all reports, and trends can be identified.

- This methodology steers away from aggregating lower-level scores to give an overall score in order to prevent issues from getting 'lost' in the aggregation process. Rather, different issues are assessed and scored at different levels (Part II and Part III). The overall score is still qualitatively informed by the process of completing the analysis of the five core elements, but without losing the unique 'story' of each evaluation report.

Special attention to Equity

Since 2011, the Evaluation function has been paying special attention to ensure that evaluations managed by UNICEF are equity-focused. In addition to the development of a specific manual on how to design and manage Equity-focused evaluation[1] (available in Arabic, French, English, Spanish and Russian), GEROS also included a specific session on Equity.

---

[1] Please visit the following website to download it: http://www.mymande.org/content/how-design-and-manage-equity-focused-evaluations

The Test

As a means of achieving consistency across a range of different reviewers, each considering a unique evaluation report, a 'test' is applied when considering what constitutes good quality. This is a practice used in some common law systems in order to enable adaptation to changing contexts (e.g. there is a test to determine if a person owes another a Duty of Care). As a body of experience is built up, the consistency with which this test can be applied also grows. The test itself can also be adapted according to experience. This fits with the peer-based quality assurance system annotated in this document.

The UNICEF GEROS gives guidance on 'good quality', noting that a good quality report has the features of being credible, addressing the evaluation questions, based on evidence, and, critically, should be able to be used with confidence.

There remain some variables within this description, such as 'who' must find the report credible, and 'what' they must be confident in using it for. Therefore, when scoring each of the five core elements and the overall report, it is proposed that the reviewers apply the following test (informed by the process of answering questions on the report):

> A report, or element of a report, is good quality if:
> - A reasonable person at a managerial grade
>
> - Existing in the context to which the evaluation speaks
>
> - Would have the confidence to act (including the allocation of funding)
>
> - Based on the Findings and Conclusions and
>
> - In pursuance of the recommendations made in the report,
>
> - In a manner that they believe to be responsible
>
> - Based upon what is written in the evaluation report.

The application of this test enables the reviewer to take into consideration the standards of partner institutions and OECD-DAC norms when reviewing joint evaluations and country-led evaluations, where this is appropriate.

Annex I: Assessment and rating matrix

Part I – the Evaluation Typology

Basic information

Title of the Evaluation Report	
Report sequence number	2010/XXX
Date of review	DD MM YYYY
Year of the Evaluation Report	YYYY
Region	List
Country(ies)	List
Type of Report	List *(if it is not an evaluation then the reviewer is not meant to review the report)*
ToRs Sent with Report	Yes or No
Name of reviewer (*Company*)	
Name of report author (company or individual)	

Classification of the Evaluation

1	Geographical *Coverage of the programme being evaluated and generalizability of evaluation findings*	**1.1 Sub-national:** The programme and evaluation covers selected sub-national units (districts, provinces, states, etc.) within a country, where results **cannot** be generalized to the whole country **1.2 National:** The programme covers the whole country, and the evaluation draws a sample in every district, or uses a sampling frame that is representative of the whole country. **1.3 Multi-country:** Where one programme is implemented in several countries, or different programmes of a similar theme are implemented in several countries, the evaluation would cover two or more countries within one region. The results of the evaluation would not be generalizable to other countries in the region. **1.4 Regional:** Where one programme is implemented in several countries, or different programmes of a similar theme are implemented in several countries, the evaluation covers multiple countries within the region and the sampling is adequate to make the results generalizable to the region. **1.5 Multi-region/Global:** The programme is implemented in two or more regions, or deliberately targets all regions. The evaluation would typically sample several countries across multiple regions, with the results intended to be generalizable in two or more regions.
2	Management of Evaluation *Managerial control and oversight of evaluation decisions (i.e., TORs, selection of consultants,*	**2.1 UNICEF managed:** Working with national partners of different categories UNICEF is responsible for all aspects of the evaluation. **2.2 Joint managed, with one or more UN agencies:** UNICEF is the co-manager with one or more UN agencies **2.3 Joint managed, with organisations outside the UN system:** UNICEF is the co-manager with one or more organizations outside the UN system **2.4 Jointly Managed with Country:** Evaluations jointly managed by the Country (Government and/or CSO) and the UNICEF CO

	budgets, quality assurance and approval of evaluation findings).In all instances, it is assumed that the management approaches include relevant national actors (e.g., government, universities, NGOs, CBOs)	**2.5 Country-led Evaluation** Evaluations managed by the Country (Government and/or CSO**)** **2.6 UNDAF Evaluation** They are a joint UN review, conducted with national partners, about the overall results expected from the UN cooperation in the country. 2.7 Not clear from report
3	Purpose *Speaks to the overarching goal for conducting the evaluation; its raison d'etre*	**3.1 Pilot:** Where a new solution, approach, or programme is being tested at a national or sub-national level, the evaluation examines the efficacy of such an intervention with the intention to determine suitability for scaling-up. **3.2 At scale:** The evaluation examines the efficacy of a programme that is being implemented at or near its maximum intended extent, with the intention of providing feedback on efficiency and the overall effectiveness of the programme to scale up focus for lessons learned. **3.3 Real-time-evaluation:** In the context of an emergency, an evaluation of the efficacy of the response, which collates lessons that can be applied back to an on-going response **3.4 Humanitarian:** Humanitarian evaluation assesses organizational performance in emergency settings (including both natural disasters and conflicts) at various phases of these crises, from preparedness and risk reduction to response, recovery and the transition to development **3.5 Project:** An evaluation which is step-by-step process of collecting, recording and organisation information about the project results including immediate results, short-term outputs and long-term project outcomes **3.6 Programme**: An evaluation of a sectorial programme to determine its overall effectiveness and efficiency in relation to the stated goals and objectives **3.7 Country Programme Evaluation (CPE):** An evaluation that assess the relevance, effectiveness, efficiency, sustainability of the entire UNICEF Country Programme **3.8 Policy:** An evaluation whose main purpose is to examine the results of a policy that is delinked from field-based programming operations. **3.9 Regional/ Multi-country programme evaluation:** An evaluation that assesses several programmes from a regional or multi-country perspective
4	Result *Level of*	**4.1 Activities and products:** Describes things that have been done rather than their effects (workshops given, publications produced, meetings attended or organized)

	changes sought, as defined in results based management: refer to substantial use of highest level reached	**4.2 Output:** Causal effects deriving directly from programme activities, and assumed to be completely under programme control **4.3 Outcome:** Effects from one or more programmes being implemented by multiple actors (UNICEF and others), where the cumulative effect of outputs elicits results beyond the control of any one agency or programme **4.4 Impact:** Final results of a programme or policy on the intended beneficiaries and, where possible, on comparison groups. Reflects the cumulative effect of donor supported programmes of cooperation and national policy initiatives.
5	MTSP Correspondence *Alignment with MTSP focus area priorities: (1) Young child survival and development; (2) Basic education and gender equality; (3) HIV/AIDS and children; (4) Child protection from violence, exploitation and abuse; and (5) Policy advocacy and partnerships for children's rights*	**5.1 Sectoral:** Addresses issues within only one of the five MTSP focus areas: (1) Young child survival and development; (2) Basic education and gender equality; (3) HIV/AIDS and children; (4) Child protection from violence, exploitation and abuse; and (5) Policy advocacy and partnerships for children's rights **5.2 Multi-sectoral:** Addresses issues in two or more MTSP focus areas **5.3 Cross-cutting:** Addresses issues that are named as cross-cutting strategies of the MTSP or otherwise known to operate within all MTSP areas. Includes but is not limited to the equity-focused and human rights-based approach to programming, gender equality, knowledge management, evaluation, and communication for development. **5.4 Organizational performance:** Evaluation of institutional effectiveness in operational areas, including supply, information technology, human resources, and finance. Also includes the more general issue of overall quality of UNICEF planning and implementation
6	Level of Independence *Implementation and control of the evaluation activities*	**6.1 Self-evaluation:** A significant component of evaluation management activities and decision-making about the evaluation are implemented by individuals associated with the target programme/intervention (e.g., programmes officer/specialists) **6.2 Independent internal:** The evaluation is implemented by consultants but managed in-house by UNICEF professionals. The overall responsibility for the evaluation lies within the division whose work is being evaluated. **6.3 Independent external:** The evaluation is implemented by external consultants and/or UNICEF Evaluation Office professionals. The overall responsibility for the evaluation lies outside the division whose work is being evaluated.
7	Stage	**7.1 Formative:** An evaluation with the purpose and aim of improving the programme. Formative evaluations strengthen or improve the object being evaluated by examining the delivery of the programme **7.2 Summative:** An evaluation that examines the effects or outcomes of the object being evaluated and summarize it by describing what happened subsequent to delivery of the programme **7.3 Summative and formative:**

		An evaluation that combines the elements of a formative and a summative evaluation.

Classification of the colour coding

Colour Coding	CC	Dark green	Green	Amber	Red	White
	Questions	Outstanding	Yes	Mostly Satisfactory	Unsatisfactory	Not Applicable
	Section & Overall Rating	Outstanding, best practice	Highly Satisfactory	Mostly Satisfactory	Unsatisfactory	

Classification of Key Questions and Cornerstone Questions

The key questions are highlighted as shown here, and are important questions in guiding the analysis of the section
The Cornerstone questions are in column J and are questions that need to be answered for rating and justification of each of the six sections

Part II – the Six Core Elements

Section A: Object of the Evaluation

Question	Rating	Remarks	A/ Does the report present a clear & full description of the 'object' of the evaluation?		Constructive feedback for future reports
	(colour coding)		*The report should describe the object of the evaluation including the results chain, meaning the 'theory of change' that underlies the programme being evaluated. This theory of change includes what the programme was meant to achieve and the pathway (chain of results) through which it was expected to achieve this.* *The context of key social, political, economic, demographic, and institutional factors that have a direct bearing on the object should be described. For example, the partner government's strategies and priorities, international, regional or country development goals, strategies and frameworks, the concerned agency's corporate goals & priorities, as appropriate.*		*Including how to address weaknesses and maintaining good practice*
1 Is the object of the evaluation well described? This needs to include a clear description of the interventions (project, programme, policies, otherwise) to be evaluated including how the designer thought that it would address the problem identified, implementing modalities, other parameters including costs, relative importance in the organization and (number of) people reached.	Outstanding Yes Mostly No N/A		• Outstanding best practice • Highly satisfactory • Mostly satisfactory • Unsatisfactory		
2 Is the context explained and related to the object that is to be evaluated? The context includes factors that have a direct bearing on the object of the evaluation: social, political, economic, demographic, and institutional. These factors may include strategies, policies, goals, frameworks & priorities at the: international level; national Government level; individual agency level	Outstanding Yes Mostly No N/A				
3 Does this illuminate findings? The context should ideally be linked to the findings so that it is clear how the wider situation may have influenced the outcomes observed.	Outstanding Yes Mostly No				

Question	Rating	Remarks	A/ Does the report present a clear & full description of the 'object' of the evaluation?		Constructive feedback for future reports
	• N/A				
4 Is the results chain or logic well articulated? The report should identify how the designers of the evaluated object thought that it would address the problem that they had identified. This can include a results chain or other logic models such as theory of change. It can include inputs, outputs and outcomes, it may also include impacts. The models need to be clearly described and explained.	• Outstanding Yes Mostly • No • N/A				
5 Are key stakeholders clearly identified? These include o implementing agency(ies) o development partners o rights holders o primary duty bearers o secondary duty bearers	• Outstanding Yes Mostly • No • N/A				
6 Are key stakeholders' contributions described? This can involve financial or other contributions and should be specific. If joint program also specify UNICEF contribution, but if basket funding question is not applicable	• Outstanding Yes Mostly • No • N/A				
7 Are UNICEF contributions described? This can involve financial or other contributions and should be specific	• Outstanding Yes Mostly • No • N/A				
8 Is the implementation status described? This includes the phase of implementation and significant changes that have happened to plans, strategies, performance frameworks, etc. that have occurred - including the implications of these changes.	• Outstanding Yes Mostly • No • N/A				
Executive Feedback on Section A Issues for this section relevant for feedback to senior management (positives & negatives), & justify rating. *Up to two sentences*					

Section B: *Evaluation Purpose, Objectives and Scope*

Question	Rating	Remarks	B/ Are the evaluation's purpose, objectives and scope sufficiently clear to guide the evaluation?		Constructive feedback for future reports
	(colour coding)		*The purpose of the evaluation should be clearly defined, including why the evaluation was needed at that point in time, who needed the information, what information is needed, and how the information will be used. The report should provide a clear explanation of the evaluation objectives and scope including main evaluation questions and describes and justifies what the evaluation did and did not cover. The report should describe and provide an explanation of the chosen evaluation criteria, performance standards, or other criteria used by the evaluators.*		*Including how to address weaknesses and maintaining good practice*
9 Is the purpose of the evaluation clear? This includes why the evaluation is needed at this time, who needs the information, what information is needed, how the information will be used.	Outstanding / Yes / Mostly / No / N/A		• Outstanding best practice • Highly Satisfactory • Mostly Satisfactory • Unsatisfactory		
10 Are the objectives and scope of the evaluation clear and realistic? This includes: Objectives should be clear and explain what the evaluation is seeking to achieve; Scope should clearly describe and justify what the evaluation will and will not cover; Evaluation questions may optionally be included to add additional details	Outstanding / Yes / Mostly / No / N/A				
11 Do the objective and scope relate to the purpose? The reasons for holding the evaluation at this time in the project cycle (purpose) should link logically with the specific objectives the evaluation seeks to achieve and the boundaries chosen for the evaluation (scope)	Outstanding / Yes / Mostly / No / N/A				

Question	Rating	Remarks	B/ Are the evaluation's purpose, objectives and scope sufficiently clear to guide the evaluation?		Constructive feedback for future reports
12 Does the evaluation provide a relevant list of evaluation criteria that are explicitly justified as appropriate for the Purpose? It is imperative to make the basis of the value judgements used in the evaluation transparent if it is to be understood and convincing. UNEG evaluation standards refer to the OECD/DAC criteria, but other criteria can be used such as Human rights and humanitarian criteria and standards (e.g. SPHERE Standards) but this needs justification. Not all OECD/DAC criteria are relevant to all evaluation objectives and scopes. The TOR may set the criteria to be used, but these should be (re)confirmed by the evaluator. Standard OECD DAC Criteria include: Relevance; Effectiveness; Efficiency; Sustainability; Impact Additional humanitarian criteria include; Coverage; Coordination; Coherence; Protection *(This is an extremely important question to UNICEF)*	• Outstanding • Yes • Mostly • No • N/A				
13 Does the evaluation explain why the evaluation criteria were chosen and/or any standard DAC evaluation criteria (above) rejected? The rationale for using each particular criterion and rejecting any standard OECD-DAC criteria (where they would be applicable) should be explained in the report.	• Outstanding • Yes • Mostly • No • N/A				
Executive Feedback on Section B Issues for this section relevant for feedback to senior management (positives & negatives), & justify rating. *Up to two sentences*					

Section C: Evaluation Methodology, Gender and Human Rights

Question	Rating	Remarks	C/ Is the methodology appropriate & sound?		Constructive feedback for future reports
	(colour coding)		*The report should present a transparent description of the methodology applied to the evaluation that clearly explains how the evaluation was specifically designed to address the evaluation criteria, yield answers to the evaluation questions and achieve the evaluation purposes.* *The report should also present a sufficiently detailed description of methodology in which methodological choices are made explicit and justified and in which limitations of methodology applied are included. The report should give the elements to assess the appropriateness of the methodology. Methods as such are not 'good' or 'bad', they are only so in relation to what one tries to get to know as part of an evaluation. Thus this standard assesses the suitability of the methods selected for the specifics of the evaluation concerned, assessing if the methodology is suitable to the subject matter and the information collected are sufficient to meet the evaluation objectives.*		*Including how to address weaknesses and maintaining good practice*
14 Does the report specify data collection methods, analysis methods, sampling methods and benchmarks? This should include the rationale for selecting methods and their limitations based on commonly accepted best practice.	• Outstanding • Yes • Mostly • No • N/A		• Outstanding best practice • Highly satisfactory • Mostly satisfactory • Unsatisfactory		
15 Does the report specify data sources, the rationale for their selection, and their limitations? This should include a discussion of how the mix of data sources was used to obtain a diversity of perspectives, ensure accuracy & overcome data limits	• Outstanding • Yes • Mostly • No • N/A				
16 Are ethical issues and considerations described? The design of the evaluation should contemplate: How ethical the initial design of the programme was; The balance of costs and benefits to participants (including possible negative impact) in the programme and in the evaluation; The ethics of	• Outstanding • Yes • Mostly • No				

Question	Rating	Remarks	C/ Is the methodology appropriate & sound?		Constructive feedback for future reports
who is included and excluded in the evaluation and how this is done	• N/A				
17 Does the report refer to ethical safeguards appropriate for the issues described? When the topic of an evaluation is contentious, there is a heightened need to protect those participating. These should be guided by the UNICEF Evaluation Office Technical Note and include: protection of confidentiality; protection of rights; protection of dignity and welfare of people (especially children); Informed consent; Feedback to participants; Mechanisms for shaping the behaviour of evaluators and data collectors	• Outstanding • Yes • Mostly • No • N/A				
18 Is the capability and robustness of the evaluated object's monitoring system adequately assessed? The evaluation should consider the details and overall functioning of the management system in relation to results: from the M&E system design, through individual tools, to the use of data in management decision making.	• Outstanding • Yes • Mostly • No • N/A				
19 Does the evaluation make appropriate use of the M&E framework of the evaluated object? In addition to articulating the logic model (results chain) used by the programme, the evaluation should make use of the object's logframe or other results framework to guide the assessment. The results framework indicates how the programme design team expected to assess effectiveness, and it forms the guiding structure for the management of implementation.	• Outstanding • Yes • Mostly • No • N/A				
20 Did the evaluation design and style consider incorporation of the UN and UNICEF's commitment to a human rights-based approach to programming, to gender equality, and to equity? This could be done in a variety of ways including: use of a rights-based framework, use of CRC, CCC, CEDAW and other rights related benchmarks, analysis of right holders and duty bearers and focus on aspects of equity, social exclusion and gender. Style includes: using human-rights language; gender-sensitive and child-sensitive writing; disaggregating data by gender, age and disability groups; disaggregating data by	• Outstanding • Yes • Mostly • No • N/A				

Question	Rating	Remarks	C/ Is the methodology appropriate & sound?	Constructive feedback for future reports
socially excluded groups.				
21 Does the evaluation assess the extent to which the implementation of the evaluated object was monitored through human rights (inc. gender & child rights) frameworks? UNICEF commits to go beyond monitoring the achievement of desirable outcomes, and to ensure that these are achieved through morally acceptable processes. The evaluation should consider whether the programme was managed and adjusted according to human rights and gender monitoring of processes.	• Outstanding • Yes • Mostly • No • N/A			
22 Do the methodology, analytical framework, findings, conclusions, recommendations & lessons provide appropriate information on HUMAN RIGHTS (inc. women & child rights)? The inclusion of human rights frameworks in the evaluation methodology should continue to cascade down the evaluation report and be obvious in the data analysis, findings, conclusions, any recommendations and any lessons learned. If identified in the scope the methodology should be capable of assessing the level of: Identification of the human rights claims of rights-holders and the corresponding human rights obligations of duty-bearers, as well as the immediate underlying & structural causes of the non realisation of rights.; Capacity development of rights-holders to claim rights, and duty-bearers to fulfil obligations.	• Outstanding • Yes • Mostly • No • N/A			

Question	Rating	Remarks	C/ Is the methodology appropriate & sound?		Constructive feedback for future reports
23 Do the methodology, analytical framework, findings, conclusions, recommendations & lessons provide appropriate information on GENDER EQUALITY AND WOMEN'S EMPOWERMENT? The inclusion of gender equality frameworks in the evaluation methodology should continue to cascade down the evaluation report and be obvious in the data analysis, findings, conclusions, any recommendations and any lessons learned. If identified in the scope the methodology should be capable of assessing the immediate underlying & structural causes of social exclusion; and capacity development of women to claim rights, and duty-bearers to fulfill their equality obligations.	• Outstanding • Yes • Mostly • No • N/A				
24 Do the methodology, analytical framework, findings, conclusions, recommendations & lessons provide appropriate information on EQUITY? The inclusion of equity considerations in the evaluation methodology should continue to cascade down the evaluation report and be obvious in the data analysis, findings, conclusions, any recommendations and any lessons learned. If identified in the scope the methodology should be capable of assessing the capacity development of rights-holders to claim rights, and duty-bearers to fulfill obligations & aspects of equity.	• Outstanding • Yes • Mostly • No • N/A				
25 Are the levels and activities of stakeholder consultation described? This goes beyond just using stakeholders as sources of information and includes the degree of participation in the evaluation itself. The report should include the rationale for selecting this level of participation. Roles for participation might include: o Liaison o Technical advisory o Observer o Active decision making The reviewer should look for the soundness of the description and rationale for the degree of participation rather than the level of participation itself.	• Outstanding • Yes • Mostly • No • N/A				

Question	Rating	Remarks	C/ Is the methodology appropriate & sound?	Constructive feedback for future reports
26 Are the levels of participation appropriate for the task in hand? The breadth & degree of stakeholder participation feasible in evaluation activities will depend partly on the kind of participation achieved in the evaluated object. The reviewer should note here whether a higher degree of participation may have been feasible & preferable.	• Outstanding • Yes • Mostly • No • N/A			
27 Is there an attempt to construct a counterfactual or address issues of contribution/attribution? The counterfactual can be constructed in several ways which can be more or less rigorous. It can be done by contacting eligible beneficiaries that were not reached by the programme, or a theoretical counterfactual based on historical trends, or it can also be a comparison group.	• Outstanding • Yes • Mostly • No • N/A			
28 Does the methodology answer the evaluation questions in the context of the evaluation? The methodology should link back to the Purpose and be capable of providing answers to the evaluation questions.	• Outstanding • Yes • Mostly • No • N/A			
29 Are methodological limitations acceptable for the task in hand? Limitations must be specifically recognised and appropriate efforts taken to control bias. This includes the use of triangulation, and the use of robust data collection tools (interview protocols, observation tools etc). Bias limitations can be addressed in three main areas: Bias inherent in the sources of data; Bias introduced through the methods of data collection; Bias that colours the interpretation of findings	• Outstanding • Yes • Mostly • No • N/A			
Executive Feedback on Section C Issues for this section relevant for feedback to senior management (positives & negatives), & justify rating. *Up to two sentences*				

Section D: Findings and Conclusions

Question	Rating	Remarks	D/ Are the findings and conclusions, clearly presented, relevant and based on evidence & sound analysis?		Constructive feedback for future reports
	(colour coding)		*Findings should respond directly to the evaluation criteria and questions detailed in the scope and objectives section of the report. They should be based on evidence derived from data collection and analysis methods described in the methodology section of the report.* *Conclusions should present reasonable judgments based on findings and substantiated by evidence, providing insights pertinent to the object and purpose of the evaluation.*		*Including how to address weaknesses and maintaining good practice*
30 Are findings clearly presented and based on the objective use of the reported evidence? Findings regarding the inputs for the completion of activities or process achievements should be distinguished clearly from results. Findings on results should clearly distinguish outputs, outcomes and impacts (where appropriate). Findings must demonstrate full marshalling and objective use of the evidence generated by the evaluation data collection. Findings should also tell the 'whole story' of the evidence and avoid bias.	Outstanding Yes Mostly No N/A		• Outstanding best practice • Highly satisfactoy • Mostly satisfactory • Unsatisfactory		
31 Do the findings address all of the evaluation's stated criteria and questions? The findings should seek to systematically address all of the evaluation questions according to the evaluation framework articulated in the report.	Outstanding Yes Mostly No N/A				
32 Do findings demonstrate the progression to results based on the evidence reported? There should be a logical chain developed by the findings, which shows the progression (or lack of) from implementation to results.	Outstanding Yes Mostly No N/A				
33 Are gaps and limitations discussed? The data may be inadequate to answer all the evaluation questions as satisfactorily as intended, in this case the limitations should be clearly presented and discussed. Caveats should be included to guide the reader on how to interpret the findings. Any gaps in the programme or unintended effects	Outstanding Yes Mostly No N/A				

Question	Rating	Remarks	D/ Are the findings and conclusions, clearly presented, relevant and based on evidence & sound analysis?		Constructive feedback for future reports
should also be addressed					
34 Are unexpected findings discussed? If the data reveals (or suggests) unusual or unexpected issues, these should be highlighted and discussed in terms of their implications.	Outstanding Yes Mostly No N/A				
35 Is a cost analysis presented that is well grounded in the findings reported? Cost analysis is not always feasible or appropriate. If this is the case then the reasons should be explained. Otherwise the evaluation should use an appropriate scope and methodology of cost analysis to answer the following questions: o How programme costs compare to other similar programmes or standards o Most efficient way to get expected results o Cost implications of scaling up or down o Cost implications for replicating in a different context o Is the programme worth doing from a cost perspective o Costs and the sustainability of the programme.	Outstanding Yes Mostly No N/A				
36 Does the evaluation make a fair and reasonable attempt to assign contribution for results to identified stakeholders? For results attributed to the programme, the result should be mapped as accurately as possible to the inputs of different stakeholders.	Outstanding Yes Mostly No N/A				
37 Are causal reasons for accomplishments and failures identified as much as possible? These should be concise and usable. They should be based on the evidence and be theoretically robust. *(This is an extremely important question to UNICEF)*	Outstanding Yes Mostly No N/A				
38 Are the future implications of continuing constraints discussed? The implications can be, for example, in terms of the cost of the programme, ability to deliver results, reputational risk, and breach of human rights obligations.	Outstanding Yes Mostly No N/A				
39 Do the conclusions present both the strengths and weaknesses of the evaluated object?	Outstanding Yes				

Question	Rating	Remarks	D/ Are the findings and conclusions, clearly presented, relevant and based on evidence & sound analysis?		Constructive feedback for future reports
Conclusions should give a balanced view of both the stronger aspects and weaker aspects of the evaluated object with reference to the evaluation criteria and human rights based approach.	• Mostly • No • N/A				
40 Do the conclusions represent actual insights into important issues that add value to the findings? Conclusions should go beyond findings and identify important underlying problems and/or priority issues. Simple conclusions that are already well known do not add value and should be avoided.	• Outstanding • Yes • Mostly • No • N/A				
41 Do conclusions take due account of the views of a diverse cross-section of stakeholders? As well as being logically derived from findings, conclusions should seek to represent the range of views encountered in the evaluation, and not simply reflect the bias of the individual evaluator. Carrying these diverse views through to the presentation of conclusions (considered here) is only possible if the methodology has gathered and analyzed information from a broad range of stakeholders.	• Outstanding • Yes • Mostly • No • N/A				
42 Are the conclusions pitched at a level that is relevant to the end users of the evaluation? Conclusions should speak to the evaluation participants, stakeholders and users. These may cover a wide range of groups and conclusions should thus be stated clearly and accessibly: adding value and understanding to the report (for example, some stakeholders may not understand the methodology or findings, but the conclusions should clarify what these findings mean to them in the context of the programme).	• Outstanding • Yes • Mostly • No • N/A				
Executive Feedback on Section D Issues for this section relevant for feedback to senior management (positives & negatives), & justify rating. *Up to two sentences*					

Section E: Recommendations and Lessons Learned

Question	Rating	Remarks	E/ Are the recommendations and lessons learned relevant and actionable?	Constructive feedback for future reports
	(colour coding)		*Recommendations should be relevant and actionable to the object and purpose of the evaluation, be supported by evidence and conclusions, and be developed with involvement of relevant stakeholders. Recommendations should clearly identify the target group for each recommendation, be clearly stated with priorities for action, be actionable and reflect an understanding of the commissioning organization and potential constraints to follow up.*	*Including how to address weaknesses and maintaining good practice*
43 Are the recommendations well-grounded in the evidence and conclusions reported? Recommendations should be logically based in findings and conclusions of the report.	• Outstanding • Yes • Mostly • No • N/A		• Outstanding best practice • Highly satisfactory • Mostly satisfactory • Unsatisfactory	
44 Are recommendations relevant to the object and the purpose of the evaluation? Recommendations should be relevant to the evaluated object	• Outstanding • Yes • Mostly • No • N/A			
45 Are recommendations clearly stated and prioritised? If the recommendations are few in number (up to 5) then this can also be considered to be prioritised. Recommendations that are over-specific or represent a long list of items are not of as much value to managers. Where there is a long list of recommendations, the most important should be ordered in priority.	• Outstanding • Yes • Mostly • No • N/A			
46 Does each recommendation clearly identify the target group for action? Recommendations should provide clear and relevant suggestions for action linked to the stakeholders who might put that recommendation into action. This ensures that the evaluators have a good understanding of the programme dynamics and that recommendations are realistic.	• Outstanding • Yes • Mostly • No • N/A			

Question	Rating	Remarks	E/ Are the recommendations and lessons learned relevant and actionable?		Constructive feedback for future reports
47 Are the recommendations realistic in the context of the evaluation? This includes: o an understanding of the commissioning organisation o awareness of the implementation constraints o an understanding of the follow-up processes	• Outstanding • Yes • Mostly • No • N/A				
48 Does the report describe the process followed in developing the recommendations? The preparation of recommendations needs to suit the evaluation process. Participation by stakeholders in the development of recommendations is strongly encouraged to increase ownership and utility.	• Outstanding • Yes • Mostly • No • N/A				
49 Are lessons learned correctly identified? Lessons learned are contributions to general knowledge. They may refine or add to commonly accepted understanding, but should not be merely a repetition of common knowledge. Findings and conclusions specific to the evaluated object are not lessons learned.	• Outstanding • Yes • Mostly • No • N/A				
50 Are lessons learned generalised to indicate what wider relevance they may have? Correctly identified lessons learned should include an analysis of how they can be applied to contexts and situations outside of the evaluated object.	• Outstanding • Yes • Mostly • No • N/A				
Executive Feedback on Section E Issues for this section relevant for feedback to senior management (positives & negatives), & justify rating. *Up to two sentences*					

Section F: The Report is well structured, logical and clear

Question	Rating	Remarks	F/ Overall, do all these elements come together in a well structured, logical, clear and complete report?		Constructive feedback for future reports
	(colour coding)		*The report should be logically structured with clarity and coherence (e.g. background and objectives are presented before findings, and findings are presented before conclusions and recommendations). It should read well and be focused.*		*Including how to address weaknesses and maintaining good practice*
51 Do the opening pages contain all the basic elements? Basic elements include all of: Name of the evaluated object; Timeframe of the evaluation and date of the report; Locations of the evaluated object; Names and/or organisations of evaluators; Name of the organisation commissioning the evaluation; Table of contents including tables, graphs, figures and annex; List of acronyms	Outstanding Yes Mostly No N/A		• Outstanding best practice • Highly satisfactory • Mostly satisfactory • Unsatisfactory		
52 Is the report logically structured? Context, purpose, methodology and findings logically structured. Findings would normally come before conclusions, recommendations & lessons learnt	Outstanding Yes Mostly No N/A				
53 Do the annexes contain appropriate elements? Appropriate elements may include: ToRs; List of interviewees and site visits; List of documentary evidence; Details on methodology; Data collection instruments; Information about the evaluators; Copy of the evaluation matrix; Copy of the Results chain. Where they add value to the report	Outstanding Yes Mostly No N/A				
54 Do the annexes increase the usefulness and credibility of the report?	Outstanding Yes Mostly No N/A				
55 Is an executive summary included as part of the report? If the answer is No, question 56 to 58 should be N/A	Outstanding Yes Mostly No N/A				

Question	Rating	Remarks	F/ Overall, do all these elements come together in a well structured, logical, clear and complete report?		Constructive feedback for future reports
56 Does the executive summary contain all the necessary elements? Necessary elements include all of: Overview of the evaluated object; Evaluation objectives and intended audience; Evaluation methodology; Most important findings and conclusions; Main recommendations	• Outstanding • Yes • Mostly • No • N/A				
57 Can the executive summary stand alone? It should not require reference to the rest of the report documents and should not introduce new information or arguments	• Outstanding • Yes • Mostly • No • N/A				
58 Can the executive summary inform decision making? It should be short (ideally 2-3 pages), and increase the utility for decision makers by highlight key priorities.	• Outstanding • Yes • Mostly • No • N/A				
Executive Feedback on Section F Issues for this section relevant for feedback to senior management (positives & negatives), & justify rating. *Up to two sentences*					

Additional Information

Question	Remarks
i/ Does the evaluation successfully address the Terms of Reference? If the report does not include a TOR then a recommendation should be given to ensure that all evaluations include the TOR in the future. Some evaluations may be flawed because the TORs are inappropriate, too little time etc. Or, they may succeed despite inadequate TORs. This should be noted under vii in the next section	
ii/ Identify aspects of good practice in the evaluation In terms of evaluation	
iii/ Identify aspects of good practice of the evaluation In terms of programmatic, sector specific, thematic expertise	

Part III – the Overall Rating

Taking into account – as appropriate - the 'type' of evaluation (refer back to Typology) and thereby the target audience for utility of the evaluation consider each of the questions below

Question	Rating (colour coding)	Remarks	OVERALL RATING Informed by the answers above, apply the reasonable person test to answer the following question: Ω/ Is this a credible report that addresses the evaluation purpose and objectives based on evidence, and that can therefore be used with confidence? This question should be considered from the perspective of UNICEF strategic management.	
i/ To what extent does each of the six sections of the evaluation provide sufficient credibility to give the reasonable person confidence to act? Taken on their own, could a reasonable person have confidence in each of the five core evaluation elements separately? It is particularly important to consider: o Is the report methodologically appropriate? o Is the evidence sufficient, robust and authoritative? o Do the analysis, findings, conclusions and recommendations hold together?	• Outstanding • Yes • Mostly • No • N/A		• Outstanding best practice • Highly satisfactory • Mostly satisfactory • Unsatisfactory	
ii/ To what extent do the six sections hold together in a logically consistent way that provides common threads throughout the report? The report should hold together not just as individually appropriately elements, but as a consistent and logical 'whole'.	• Outstanding • Yes • Mostly • No • N/A			
iii/ Are there any reasons of note that might explain the overall performance or particular aspects of this evaluation report? This is a chance to note mitigating factors and/or crucial issues apparent in the review of the report.				
Executive Feedback on Overall Rating Issues for this section relevant for feedback to senior management (positives & negatives), & justify rating. *Up to two sentences*				

Executive Review Format

This is the proposed format for review feedback to be presented to international and regional decision makers.

Please note that this template will automatically update while the Review template is being completed

Colour Coding	CC	Dark green	Green	Amber	Red
	Section & Overall Rating	Outstanding, best practice	Highly Satisfactory	Mostly Satisfactory	Unsatisfactory

	Response			
Title of the Evaluation Report	**0**			
Report sequence number		**Date of Review**	**Year of the Evaluation Report**	**0**
Region	0		**Country(ies)**	0
Executive Summary in Final Report	0		**TORs sent with Report**	0
OVERALL RATING			**Overall**	**0**
Executive Feedback on Overall Rating	0			
SECTION A: OBJECT OF THE EVALUATION			**Section A Colour Coding**	**0**
Executive Feedback on Section A	0			
SECTION B: EVALUATION PURPOSE, OBJECTIVES AND SCOPE			**Section B Colour Coding**	**0**
Executive Feedback on Section B	0			
SECTION C: EVALUATION METHODOLOGY, GENDER, HUMAN RIGHTS			**Section C Colour Coding**	**0**
Executive Feedback on Section C	0			
SECTION D: FINDINGS AND CONCLUSIONS			**Section D Colour Coding**	**0**
Executive Feedback on Section D	0			
SECTION E: RECOMMENDATIONS AND LESSONS LEARNED			**Section E Colour Coding**	**0**
Executive Feedback on Section E	0			
SECTION F: REPORT IS WELL STRUCTURED, LOGIC AND CLEAR			**Section F Colour Coding**	**0**
Executive Feedback on Section F	0			

Annex 2: Definition of evaluation and other forms of assessment/data collection and analysis as by PPP Manual

An evaluation is defined as an assessment, as systematic and impartial as possible, of an activity, project, programme, strategy, policy, topic, theme, sector, operational area and institutional performance. It focuses on expected and achieved accomplishments, examining the results chain, processes, contextual factors an causality, in order to understand achievements or the lack thereof. It aims at determining the relevance, impact, effectiveness, efficiency and sustainability of interventions.

Other forms of assessment being conducted in UNICEF vary in purpose and level of analysis, and may overlap to some extent. Evaluation is to be differentiated from the following:

a. *Needs assessment*: A process or a systematic set of procedures undertaken for the purpose of setting priorities and making decisions about program or organizational improvement or allocation of resources. The priorities come from identified needs which are measured discrepancies (gaps) between the current (what is) state of affairs of a group or organization and the desired (what should be) state in regard to variables of interest.

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

c. *Evaluability.* Extent to which a program can be evaluated in a reliable and credible fashion. Evaluability assessment calls for the early review of a proposed program in order to ascertain whether its objectives are adequately defined and its results verifiable.

d. *Review, including Mid-Term Review*: The periodic or *ad hoc* assessments of the performance of an undertaking, which do not apply the due process of evaluation.

e. *Inspection*: A general examination that seeks to identify vulnerable areas and malfunctions and to propose corrective action.

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

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

h. *Research and study*: A systematic examination designed to develop or contribute to knowledge. A baseline study is an analysis describing the situation prior to a development intervention, against which progress can be assessed or comparisons made.

i. *Survey:* An exercise to collect data on the situation of children and women. Major surveys supported by UNICEF are the Multiple Indicators Cluster Survey (MICS) and Demographic and Health Survey (DHS).

l. *Internal management consulting*: Consulting services to help managers implement changes which address organizational and managerial challenges and improve internal work processes.

Annex 3: UNICEF-adapted UNEG Evaluation reports standards

1. The report Structure

1.0 The report is well structured, logical, clear and complete
1.1 Report is logically structured with clarity and coherence (e.g. background and objectives are presented before findings, and findings are presented before conclusions and recommendations)

1.2 The title page and opening pages provide key basic information
1. Name of the evaluation object
2. Timeframe of the evaluation and date of the report
3. Locations (country, region, etc) of the evaluation object
4. Names and/or organizations of evaluators
5. Name of the organization commissioning the evaluation
6. Table of contents which also lists Tables, Graphs, Figures and Annexes
List of acronyms

1.3 **Executive Summary** is a stand-alone section of 2-3 pages that includes:
1. Overview of the evaluation object
2. Evaluation objectives and intended audience
3. Evaluation methodology
4. Most important findings and conclusions
5. Main recommendations
1.4 Annexes increase the credibility of the evaluation report. They may include, inter alia:
1. ToRs
2. List of persons interviewed and sites visited
3. List of documents consulted
4. More details on methodology, such as data collection instruments, including details of their reliability and validity
5. Evaluators biodata and/or justification of team composition
6. Evaluation matrix
Results framework

2. Object of Evaluation
2.0 The report presents a clear and full description of the 'object' of the evaluation

2.1 The **logical model and/or the expected results chain** (inputs, outputs, and outcomes) of the object is clearly described

2.2 The **context of key social, political, economic, demographic, and institutional factors** that have a direct bearing on the object is described. For example, the partner government's strategies and priorities, international, regional or country development goals, strategies and frameworks, the concerned agency's corporate goals and priorities, as appropriate

The scale and complexity of the object of the evaluation are clearly described , for example:
2.3 **The number of components**, if more than one, and the size of the population each component is intended to serve, either directly or indirectly

- **The geographic context and boundaries** (such as the region, country, and/or landscape and challenges where relevant).

- The purpose and goal, and organization/management of the object

- The **total resources** from all sources, including human resources and budget (s) (e.g. concerned agency, partner.

2.4 The **key stakeholders involved** in the object implementation, including the implementing agency (s) and partners, other key stakeholders and their roles

2.5 The report identifies **the implementation status of the object**, including its phase of implementation and any significant changes (e.g. plans, strategies, logical frameworks) that have occurred over time and explains the implications of those changes for the evaluation

3. Evaluation Purpose, Objective(s) and Scope
3.0 The evaluation's purpose, objectives and scope are fully explained

3.1 The purpose of the evaluation is clearly defined, including why the evaluation was needed at that point in time, who needed the information, what information is needed, how the information will be used.

3.2 The report should provide a clear explanation of the evaluation objectives and scope including main evaluation questions and describes and justifies what the evaluation did and did not cover

3.3 The report describes and provides an explanation of the chosen evaluation criteria, performance standards, or other criteria used by the evaluators

3.4 As appropriate, evaluation objectives and scope include questions that address issues of gender and human rights

4. Evaluation Methodology
4.0 The report presents transparent description of the methodology applied to the evaluation that clearly explains how the evaluation was specifically designed to address the evaluation criteria, yield answers to the evaluation questions and achieve evaluation purposes.

4.1 The report describes the data collection methods and analysis, the rationale for selecting them, and their limitations. Reference indicators and benchmarks are included where relevant.

4.2 The report describes the data sources, the rationale for their selection, and their limitations. The report includes discussion of how the mix of data sources was used to obtain a diversity of perspectives, ensure data accuracy and overcome data limits.

4.3 The report describes the sampling frame – area and population to be represented, rationale for selection, mechanics of selection, numbers selected out of potential subjects, and limitations of the sample

4.4 The evaluation report gives me complete description of stakeholder's consultation process in the evaluation including the rationale for selecting the particular the particular level and activities of consultation

4.5 The methods employed are appropriate for the evaluation and to answer its questions.

4.6 The methods employed are appropriate for analysing gender and human rights issues including child rights issues identified in the evaluation scope.

4.7 The report presents evidence that adequate measures were taken to ensure data quality, including evidence supporting the reliability and validity of data collection tools (e.g. interview protocols, observation tools etc)

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

5. Findings
4.0 Findings respond directly to the evaluation criteria and questions detailed in the scope and objectives section of the report are based on evidence derived from data collection and analysis methods described in the methodology section of the report

5.1 Reported findings reflect systematic and appropriate analysis and interpretation of the data

5.2 Reported findings address the evaluation criteria (such as efficiency, effectiveness, sustainability, impact and relevance) and questions defined in the evaluation scope

5.3 Findings are objectively reported on the evidence

5.4 Gaps and limitations in the data and/or unanticipated findings are reported and discussed

5.5 Reasons for accomplishments and failures, especially continuing constraints, were identified as much as possible

5.6 Overall findings are presented with clarity, logic and coherence

6. Conclusions and Lessons Learned
6.0 Conclusions present reasonable judgements based on findings and substantiated by evidence and provide insights pertinent to the object and purpose of the evaluation

6.1 The conclusions reflect reasonable evaluative judgements relating to key evaluation questions

6.2 Conclusions are well substantiated by the evidence presented and are logically connected to evaluation findings

6.3 Stated conclusions provide insights into the identification and/or solutions of important problems issues pertinent to the prospective decisions and actions of evaluation users

6.4 Conclusions present strengths and weaknesses of the object (policy, programmes, projects or other intervention) being evaluated, based on the evidence presented in taking due account of the views of a diverse cross-section of stakeholders

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

7. Recommendations
7.0 Recommendations are relevant to the object and purpose of the evaluation, are supported by evidence and conclusions, and were developed with involvement of relevant stakeholders

7.1 The report describes the process followed in developing the recommendation including consultation with stakeholders

7.2 Recommendations are firmly based on evidence and conclusions

7.3 Recommendations are relevant to the object and purpose of the evaluation

7.4 Recommendations clearly identify the target group for each recommendation

7.5 Recommendations are clearly stated with priorities for action made clear

7.6 Recommendations are actionable and reflect an understanding of the commissioning organization and potential constraints to follow up

8. Gender and Human rights, including child rights
8.0 The report illustrates the extent to which the design and implementation of the object, the assessment of results and the evaluation process incorporates a gender equality perspective and human rights based approach, including child rights

8.1 The report uses gender sensitive, child sensitive and human rights-based language throughout, including data disaggregated by sex, age and disability

8.2 The evaluation approach and data collection and analysis methods are gender equality, and human rights including child rights responsive and appropriate for analyzing the gender equality, human rights issues including child rights identified in the scope

8.3 The report assesses if the design of the object was based on a sound gender analysis, and human rights analysis including child rights and implementation for results was monitored through gender, and human rights frameworks including child rights, as well as the actual results on gender equality, and human rights including child rights

8.4 Reported findings, conclusions, recommendations and lessons provide adequate information on gender equality, and human rights aspect including child rights.