

GEROS Evaluation Quality Assurance Tool

Version: September 2021

Developmental evaluation of UNICEF's internal communication and staff engagement function

REPORT RATING SUMMARY			
Overall Rating	64%	Satisfactory	
●●●●●	Exceptional (96% - 100%)	5	
●●●●	Highly Satisfactory (87.5% - 95.99%)	4	
●●●●	Satisfactory (62.5% - 87.49%)	3	Meets UNICEF/UNEG standards for evaluation reports. Decision makers may use the evaluation with confidence
●●●	Fair (35% - 62.49%)	2	
●●	Unsatisfactory (0% - 34.99%)	1	
REPORT DETAILS			
Title of the evaluation report	Developmental evaluation of UNICEF's internal communication and staff engagement function		
Report sequence number	Evaluation Office 65-2024-19440		
Region	HQ		
Year of report	2024		
Office	HQ		
Coverage (countries)	None		
ToRs present	Yes		
Date of review (dd/mmm/yyyy)	December 12, 2024		
Name of review firm	IODPARC		
CLASSIFICATION OF EVALUATION REPORT			
Management of evaluation (Managerial control and oversight of evaluation)	UNICEF managed		
Unicef goal areas (Alignment with strategic plan priorities)			
Every child survives and thrives	Yes		
Every child learns	Yes		
Every child is protected from violence and exploitation	Yes		
Every child lives in a safe and clean environment	Yes		
Every child has an equitable chance in life	Yes		
Gender equality (cross-cutting)	Yes		
Humanitarian action (cross-cutting)	No		
Evaluation object	System		
Evaluation type	Summative and formative		
Evaluation strategy	Mixed methods		
Evaluation design (primary method used)	Other		
Evaluation level	Output & Outcome		
Geographic scope	Multi-region/Global		
Primary SDG(s) covered (number)	None (internally-focused)		
EQA Summary: <i>The rater will provide top line issues for this evaluation relevant for feedback to senior management (positive and negative), summarizing how the evaluation report meets or fails to meet all criteria. As relevant, the rater will highlight best practice/added value elements and the level of complexity of the evaluation.</i>			
<p>Overall this is a relatively strong and interesting evaluation of a very important element of the internal processes of UNICEF. The evaluators have adopted a non-standard and quite innovative approach to the evaluation (albeit at the prompting of UNICEF), and have made good efforts to explain both the conceptual background and the process to the evaluation. This is important given that it is both a complex approach to the evaluation and is also related to the internal dynamics of UNICEF which are not easily apparent to those that may not be deeply familiar with UNICEF and its organisational culture. The evaluators do a largely good job of addressing this, though there are places where the descriptions and analysis are dense and challenging to understand. Some more specific strengths are as follows:</p> <ul style="list-style-type: none"> •A comprehensive and detailed institutional background and evolution of ICSE since 2008. •The evaluation purpose and objectives are clearly articulated and tailored to emerging needs. •The use of an innovative developmental evaluation approach emphasizing participatory methods. •The report is logically structured and generally well-organized. •Use of multiple data sources (qualitative, quantitative, and secondary) to provide a robust evidence base. •Recommendations are well-structured, actionable, and aligned with the evaluation conclusions. •Strong analysis of ICSE as a strategic enabler, addressing the core evaluation questions systematically (although some of the later sections have more cursory analysis). •The findings effectively explore causal factors influencing ICSE success and barriers. •Recommendations include cost implications, helping guide resource allocation. 			
Recommendations for Improvement: <i>The rater will identify topline recommendations to improve the evaluation, and be specific to the sections of the report where shortcomings were found. As relevant, resources will be cited to assist evaluation managers in overseeing future evaluations.</i>			

The main shortcomings identified are as follows:

- The evaluation is somewhat HQ-centric - there was quite limited representation from country operations, and field-based staff represented a very small proportion of respondents, despite being the vast majority of UNICEF personnel. There was some evidence from the field (e.g. examples of good/promising practices) but it is quite limited.
- The introduction omits some descriptive/demographic information about UNICEF staff, e.g. gender, location.
- Findings are not presented as clearly defined, numbered statements, requiring readers to infer the findings from the summaries.
- The sampling strategy lacks detail and includes biases, particularly under-representation of country office staff and does not accurately describe the validity/representativeness of the surveys.
- Ethical considerations are briefly noted but lack specific mitigation strategies for potentially sensitive issues that have been noted in the context/background.
- Analysis of survey data and secondary data is insufficiently described.
- Limited reference to gender and equity considerations despite their relevance to ICSE.
- Some recommendations (e.g., on communication channels) are not fully aligned with the evidence.
- Findings progressively lack depth as the analysis progresses through the Learning Framework.
- Lessons learned are not systematically captured and not presented in the relevant section, despite some potentially interesting examples noted throughout the report.
- There is no description of how recommendations were developed or validated.
- Quantitative survey tools are missing from the annexes.

SECTION RATINGS			
SECTION A:	EXECUTIVE SUMMARY (weight 5%)	100%	Comments on Rating
Question 1.	Can the executive summary inform decision-making?		
	i Is clearly presented, serves as a standalone document useful for informing decision making, and is of relevant conciseness and depth for key users (Maximum of 5 pages unless otherwise specified in ToR).	Yes	The executive summary is good - it is clear and concise, but has plenty of depth and insight. It can serve as a useful standalone document for the casual reader and help to inform decision making. There is no page limit mentioned in the TOR - at six pages, the report is longer than the criterion, but not excessive given that it covers an organisation-wide evaluation.
	ii Includes all necessary elements (overview of the intervention, evaluation purpose, objectives and intended audience, evaluation methodology, key conclusions on findings, lessons learned if requested, key recommendations) as per the ToR.	Yes	All key elements of the evaluation are present, with the structure of the executive summary as per the main report. Both the conclusions and recommendations have been summarised well, with the key data around priority, timeframe, responsible units and cost implications clearly noted, which will be of good use to decision-makers reviewing the summary.
	iii Includes all significant information needed to understand the intervention and the evaluation AND does not introduce new information from what is presented in the rest of the report.	Yes	The executive summary includes a good summary of relevant information from the main report. Indeed, it improves on the absence of clearly-itemised findings as found in the main report by noting in bold the headline findings per each subheading. This makes the summary very readable and useful. It reflects on the key areas that are explored in some depth in the analysis - incorporating the summary text (in blue highlights) before each findings subsection, but augmenting it with additional detail.
SECTION B:	BACKGROUND (weight 5%)	42%	Comments on Rating
Question 2.	Is the object of the evaluation clearly described?		
	i Clear and relevant description of the intervention, including: location(s), timelines, cost/budget, and implementation status.	Partially	Section 1.1 provides ample detail on the evolution of ICSE since 2008, with descriptions of the various strategic shifts, proposal developments, strategic and budget alignments, timelines and implementation status. While important, the level of detail overall is difficult to parse. The section is greatly helped by the inclusion of a summary timeline of the main developments in Box 1. The authors might consider summarising pages 1-5 in a single page and moving much of this text to an annex.
	ii Clear and relevant description of intended rightsholders (beneficiaries) and duty bearers (state and non-state actors with responsibilities regarding the intervention) by type (i.e., institutions/organizations; communities; individuals...), by geographic location(s) (i.e., urban, rural, particular neighbourhoods, town/cities, sub-regions...) and in terms of numbers reached, with disaggregation by gender, age, disability . . . (as appropriate to the purpose of the evaluation).	No	Although this is an evaluation of UNICEF's internal organizational capacity, and as such does not deal with rights-holders in the traditional sense, UNICEF staff are the beneficiaries of a robust ICSE strategy, so can be deemed this group. As such, the evaluation does not discuss the staff of UNICEF in the introduction, and this is an omission. The introduction and context would benefit from discussing the people of UNICEF - i.e. the 17,000+ staff members, their breakdown across locations, offices, grades, disaggregating by gender, disability and other key dimensions (e.g. those identifying as LGBQT, if available). These dimensions all have a bearing on the value and robustness of UNICEF's ICSE policies and practices, so warrant some description.
Question 3.	Is the context of the intervention clearly described?		

	i Clear and relevant description of the context of the intervention (i.e. relevant policy, socio-economic, political, cultural, power/privilege, institutional, international factors) and how context relates to the implementation of the intervention.	Partially	The description of the institutional background to ICSE in UNICEF is extremely detailed overall, but is somewhat difficult to grasp due to the dense narrative, although helped by the use of summaries (boxes 1 and 2). This said, in places the background and context description is somewhat unclear (particular technical terms), and in some jarring and abrupt. For example, section 1.1 (pg. 2, first para) refers to "digital communities", presumably within UNICEF - but without explanation of what they are, what their purpose is/was etc. (in the context of UNICEF, rather than generically). Another example is the next paragraph, which states quite abruptly that the internal communications infrastructure of UNICEF "had contributed to significant harm" - some more nuance in how such analysis is presented is warranted, in particular with reference to the impact on the wider staff population of UNICEF, whose needs and rights should underpin an equitable and fit-for-purpose ICSE system.
	ii Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Not Rated	This is an evaluation of UNICEF's internal organizational capacity, and as such there is no direct bearing on specific SDGs
	iii Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	No	As noted above, the core constituency of the UNICEF ICSE policy: UNICEF staff, are not discussed in this section, and should be. There are references to increases in staff digital engagement, establishment of "digital communities" etc. but little explanation or unpacking of what these imply around the needs and status of the many UNICEF staff that are affected by the ICSE policies. In particular, the (somewhat abrupt) reference to the "harm" contributed to by the pre-existing policies could and should warrant unpacking - the prevalence of SEAH, discrimination etc. that are noted on page 2 should be further explored, as they significantly underpin the argument for an improved ICSE system.
Question 4.	Are key stakeholders, their relationships and contributions clearly identified?		
	i Identification of implementing agency(ies), development partners, right holders, and additional duty bearers and other stakeholders; and of linkages between them (e.g., stakeholder map) (if relevant).	Partially	Insofar as this is a fully internal evaluation, it focuses on UNICEF business units rather than external 'duty-bearers' (other than the participation of external service providers such as Deloitte). In this, the background does a reasonable job of describing the different organizational stakeholders that have been involved in the ICSE process (if confusing to an outsider - but as this evaluation is destined for an internal audience, it is expected that the units and their relationships will be clearer).
	ii Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Yes	The background section goes into considerable detail around the description of the contribution of various internal stakeholders to the ICSE process - perhaps overly detailed (as noted above). The summary of the process and outputs of different reviews in boxes 1 and 2 are particularly useful in this regard.
SECTION C: EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)		80%	Comments on Rating
Question 5.	Is the purpose of the evaluation clearly described?		
	i Purpose of evaluation is clearly defined, including why it was needed at that point in time, its intended use, and key intended users.	Yes	The purpose of the evaluation is clear - the background and context provide ample detail on the justification and need for such an evaluation, and the purpose section (1.2) explains the rationale clearly. It also notes how the purpose as originally articulated in the TOR has been tailored to the emerging needs and dynamics of ICSE and the UNICEF strategic plan. The section clearly notes the audience for the evaluation (primary stakeholder business units within UNICEF).
Question 6.	Are the objectives and scope of the evaluation clear and realistic?		
	i Clear and complete description of what the evaluation seeks to achieve by the end of the process with reference to any changes made to the objectives included in the ToR (if applicable).	Partially	Page 9 details the objectives of the evaluation, which are clear when taken at face value. There has been a considerable degree of review and revision of the originally-stated objectives in the TOR, which is detailed in the narrative accompanying the list of objectives. The chronology for changes is quite confusingly written, however. Indeed, the final list of objectives has an omitted objective of the IR (objective 4: <i>Assess the matrix management approach as part of the operationalization of ICSE</i>) which is not referenced, either as an addition (to the originally specified objectives in the TOR) or removal (from the list in the IR). The evaluators could have made this clearer.

	ii	Clear and relevant description of the scope of the evaluation: what will and will not be covered (thematically, chronologically, geographically with key terms defined), as well as, if applicable, the reasons for this scope (e.g., specifications by the ToRs, lack of access to particular geographic areas for political or safety reasons at the time of the evaluation, lack of data/evidence on particular elements of the intervention).	Yes	The scope for the evaluation is quite broadly defined by intention, as the nature of the developmental evaluation is such that scope "is determined by what is being developed", so the evaluators are free to explore whatever elements of UNICEF operations relate to ICSE. This is reasonably well-explained in the report, although some reference to the timescale that the evaluation covers (possible reflecting back to 2019/2020 when ICSE was first mainstreamed as a strategy? Or earlier?) would have been useful. As the evaluation is internally-focused, no location considerations apply.
Question 7.		Is the theory of change, results chain or logic well articulated?		
	i	Clear description of the intervention's intended results, or of the parts of the results chain that are applicable to, or are being tested by, the evaluation.	Yes	The "learning framework" in Section 1.2 (table 1) presents the specific UNICEF Goal Areas under which ICSE is expected to contribute (these are also articulated in narrative form in section 1.1, pg. 3). It is clear that these are the results which effective ICSE will contribute to. Further, exploration of the TOC developed via participatory workshops forms a key part of the analysis under the strategic enabling function of ICSE.
	ii	Causal relationship between outputs and outcomes is presented in narrative and/or graphic form (e.g., results chain, logic model, theory of change, evaluation matrix).	Partially	Inherent to the developmental evaluation approach is a focus on the process of change, rather than a prescriptive focus on discrete outputs and outcomes, which is well explained in the preamble to the methods section (2.1) which introduces DE and how it differs from 'mainstream' evaluations. This said, the evaluators did develop a theory of change for ICSE as part of the evaluation - indeed this was a key part of the preparatory process of the evaluation, with a workshop with staff held in late 2023. The TOC is not articulated in the methods/design section, however (only a brief reference in section 2.2) - the evaluators should make clearer reference to this in the methods section and refer to the appropriate location in the report where the reader can review the TOC.
	iii	For theory-based evaluations, the theory of change or results framework is assessed, and if requested in the ToR, it is reformulated/improved by the evaluators.	Not Rated	As this is not a theory-based evaluation, this criterion is not rated.
SECTION D:		EVALUATION DESIGN AND METHODOLOGY (weight 20%)	45%	Comments on Rating
Question 8.		Does the evaluation use questions and the relevant evaluation criteria that are explicitly justified as appropriate for the purpose of the evaluation? <i>UNICEF evaluation standards refer to the OECD/DAC criteria - Relevance; Coherence; Effectiveness; Efficiency; Sustainability; Impact (not all are necessarily relevant for all evaluations). Evaluations should also consider equity and leaving no-one behind, gender and human rights based approach (these can be mainstreamed into other criteria). Humanitarian evaluations should also consider Coverage; Connectedness; Coordination; Protection; Security.</i>		
	i	Evaluation questions and sub-questions are appropriate for meeting the objectives and purpose of the evaluation. The relevant criteria are specified and are aligned with the questions.	Partially	The evaluators apply a "learning framework" approach to conceptualisation of the evaluation. This articulates a range of key questions in a grid format (cross-referenced against UNICEF goal areas) and underlying structural elements (collaboration structures, awareness/understanding, enabling environment). This framework (the text notes) was developed with the UNICEF ICSE team early in the evaluation and shared within UNICEF for feedback, so should reflect a strong consensus. The questions themselves are coherent with respect to the overall grid headings and the consensus-based approach means that they should accurately reflect the purpose and objectives of the evaluation. One issue is that the structural elements (reviewer's term) are not explained in the report - how they were derived and finalised and if others were considered. This could be easily explained with a few sentences in the narrative. Given the nature of this evaluation, formal (DAC/OECD) evaluation criteria do not apply.

	ii In addition to the questions and sub-questions, the evaluation matrix includes indicators, benchmarks, assumptions and/or other processes from which the analysis can be based and conclusions drawn.	Partially	The learning framework does present 12 key questions (and two overarching questions) that correspond roughly to a more traditional 'evaluation matrix' and should guide the collection of data via appropriately-designed tools. While this process may have been clear to the evaluators, there is no clear pathway presented in the main report or in the annexes that demonstrates how the research tools (e.g. the list of questions for interviews presented in Annex 4) are derived from these questions. The need to be focused on the process rather than overly-prescriptive and rigid frameworks is made clear, but at the same time, templates for the data collection tools were required (and developed) and the conceptual linkages between these and the learning framework is where there is a gap that should be addressed.
Question 9.	Does the report specify adequate methods for data collection, analysis, and sampling?		
	i Evaluation design and set of methods are relevant and adequately robust for the evaluation's purpose, objectives and scope; and are fully and clearly described.	Partially	The data collection and analysis section (2.2) is short (just over two pages), and while it provides a good overview of the tools and approach, it is insufficiently detailed regarding the selection, sampling and analytical approach and the justification of the approaches used. The report would greatly benefit from a dedicated appendix for the methodology, drawing heavily on the considerably more detail that is provided in the Inception Report. Given that the approach to the evaluation is somewhat 'non-standard', the evaluators might consider including some endnotes from published literature around development evaluations supporting the robustness of the chosen approach. Finally, while the interview tools are presented in the annexes, the surveys are not presented as standalone tools - the survey questions are only presented (in Annex 5) with the results under each individual question, so the tools cannot be reviewed in isolation. This is an omission.
	ii Data sources are appropriate - these would normally include qualitative and quantitative sources (unless otherwise specified in the ToR) - and are all clearly described.	Partially	The data sources are generally good - a combination of qualitative (interview) primary data, qualitative (documentary) secondary and quantitative (survey) primary data - the varied approaches represent a good mix of types of data. The sources of the data are reasonable (a mix of internal/external, HQ/region/country sources), although the precise proportions of each (specifically for the interviews) deserve some scrutiny. Specifically, the quite limited representation of staff from country offices (which constitute the vast majority of UNICEF's personnel) via only four interviews is concerning. Further, the solicitation of inputs from external perspectives is positive, but it is curious that no sister UN agencies were included which might provide comparable insights. Instead an NGO, the World Bank, a consultant firm and two service providers (that have potential conflicts of interest and thus might be definitively <u>not</u> appropriate) were selected. The basis of this is not explained.
	iii Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or, if not, provide reasons for this).	Partially	There is some, but limited, amount of information on the sampling process. While the report notes the total sample selected for each method, how individuals were selected within these is essentially purposive (although this is not specifically noted). For example, selection of participants from the four country offices included in the interviews is not explained beyond having (either) been suggested by other key informants (snowball sampling is noted as another method, but this is what suggestions from others are, so is repetitive) or been identified through the global survey, though what criteria were applied here is not explained. Similarly, the selection of country offices for participation is not explained - were the offices selected purposively, or were the individuals selected based on other criteria, irrespective of their location globally. The sampling strategy for the survey tools is similarly lacking in detail. How were the targets for the various surveys selected (e.g. were contractors/temporary contracts included)? What methods were employed to encourage responses (which appear to have been very low - less than 10% of those reached out to)? Further, the claim that the response rates were "statistically representative" is not valid - statistical representativeness is not based simply on the response rate, it is about results being generalisable to the overall population - this is determined by looking at the characteristics of the responders, which does not appear to have been done.

	iv Clear and complete description of the methods of analysis.	No	The analysis subsection is neither clear nor complete. It is a (half-page) narrative description of which a significant proportion is a discussion of the theoretical approaches to analysis using a developmental evaluation approach. While this is useful as a preamble, a more structured approach as to how this was applied to all of the collected data is largely missing. There are several sentences which note instances where data was analysed or synthesised (e.g. into learning memos) but little about how data was compiled, collated, synthesised and how themes were determined, triangulated and tested. The final sentence of the subsection does note that qualitative data was coded (using the learning framework and via open coding) but this is very brief and there is no mention of the treatment of survey data or secondary data.
	v Methodology allows for drawing causal connections between outputs and expected outcomes.	Partially	Given that different approaches were used in the methods, there is a good likelihood that causal connections could be and were drawn in line with the learning framework that underpins the evaluation. However, the shortcomings of the methods (limitations on the depth of consultation across all staff and locations, as discussed above) and the lack of important details in the methods section (also noted above), mean that a definitive answer on this criterion cannot be given.
	vi Clear and complete description of limitations and constraints faced by the evaluation, including gaps in the evidence that was generated and mitigation of bias, and how these were addressed by the evaluators (as feasible).	Partially	The limitations section (2.3) contains three relevant and well-articulated limitations/risks to the evaluation, and discusses some of the mitigations that are in place around them. These are good, but it is quite a limited list and there are likely other limitations that are not addressed. For example, the (very) limited response rate to the staff surveys introduces a significant bias. Another potential source of methodological bias is the limitations on the sample of key informants - skewed heavily towards staff at HQ or regional levels. The voices of the country-based majority of UNICEF staff may well not be represented in this evaluation.
Question 10.	Are ethical issues and considerations described? The evaluation should be guided by the UNEG ethical standards for evaluation. As such, the evaluation report should include:		
	i Explicit and contextualized reference to the UNEG obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability) and/or the principles in the 2020 revised UNEG Ethical Guidelines (integrity, accountability, respect, beneficence).	Partially	There is explicit, but not contextualised, reference to the UNEG standards (although not the 2020 version) in a very short quality standards and ethics section, with a link to UNICEF's evaluation resources page which has onward links to the suite of approaches. This is a minimally adequate approach.
	ii Description of ethical safeguards for participants appropriate for the issues relevant to methodology and how they are applied (respect for dignity and diversity, right to self-determination, fair representation, compliance with codes for vulnerable groups, confidentiality, and avoidance of harm). For those cases where the evaluation involves interviewing children, explicit reference is made to the UNICEF procedures for Ethical Research Involving Children.	No	Given that the evaluation is internal to UNICEF, and thus did not include rights-holders, the standard for ethical safeguards might be deemed to be lower than with other evaluations that engage at field level. However, as is noted in the context (pg. 2), UNICEF staff are not immune to harm, with specific issues of "gender discrimination, sexual harassment, and abuse of authority" noted from the outset of the report. Thus, an evaluation that explores staff engagement may well come into contact with instances of the above issues or at least need to consider them. Further, the interpersonal nature of management and staff engagement means that power dynamics and interpersonal relationships are a key dimension. Therefore, the need to ensure confidentiality, security, ethical use of data etc. is significant both to avoid negative repercussions for respondents or simply provide assurances to respondents that these have been comprehensively considered and mitigated. The relevant subsection provides little discussion of this. While the interview tools in the annexes do include consent-seeking text (which is positive), this text is contradictory, noting " it may be that due to your positionality and the timing of this conversation, you may be identifiable, and thus this conversation is entirely confidential. " (emphasis added). The survey tools were not included in Annex 5 (as is claimed) so any similar language could not be reviewed for these, hence this criterion is poorly met.
Question 11.	Does the evaluation incorporate innovative practice that adds value to the evaluation process?		

i Innovation practice is used to improve the quality of evaluation process. This could be evident in several ways such as the design of the methodology (i.e. use of technology for data gathering, extensive participatory processes, systematic analysis processes such as collaborative outcomes reporting and incorporation of big data, specific strategies to address complexity such as outcome harvesting, strong child rights focus), or ways of sharing of evaluation results.

Yes

The concept and design of the evaluation represent an innovative shift from more 'traditional' evaluation practice. The evaluators, based on guidance from UNICEF (as articulated in the TOR) adopted a development evaluation approach, which is highly participatory, involving embedding of evaluators within the process being evaluated (in this case within UNICEF as an organisation). The process, from the outset, appears to have been highly collaborative, with a range of communications, learning memos, observations, consultations etc. throughout. This process was integrated into both the design and rollout of the evaluation.

SECTION E: EVALUATION FINDINGS (weight 25%)		75%	Comments on Rating
Question 12.	Do the findings clearly address all evaluation objectives and scope?		
	i Findings marshal sufficient levels of evidence to systematically address all of the evaluation's questions, sub-questions and criteria.	Partially	From a systematic perspective, the findings section partially, but not completely, follows the analytical outline set in the Learning Framework (table 1). The questions articulated under column 1 (ICSE as a strategic enabler) are the most systematically and comprehensively-answered questions in the framework. As the findings progress through the framework, however some questions are combined into single analyses. For example, for GA1-3, the first and third questions are combined into one section, and the extent of the analysis is somewhat brief. In particular, for GA3, they comprise the final (short - one para) section of the findings report. This said, the evaluators do make a good effort to at least cover the questions in the framework, and there is a useful preamble analysis of progress on ICSE in UNICEF during the course of the evaluation which sets out some of the key challenges that UNICEF faces on an ongoing basis, including perceptions and understanding of the meaning of ICSE at different levels.
	ii Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Yes	The evaluators present an analysis of the theory of change that was constructed by the evaluators with the participation of a range of UNICEF stakeholders in the findings section (pg. 24). While the TOC is used in a somewhat limited manner, it does clearly link ICSE to UNICEF's results for children, and as such can be an important tool - if not in this evaluation, then maybe as an advocacy element for promotion of ICSE. The ToC was also applied to analysis of the strategic enablers of UNICEF's ICSE under question 2, linking to the UNICEF management model.
Question 13.	Are evaluation findings derived from the conscientious, explicit and judicious use of the best available, objective, reliable and valid data and by accurate quantitative and qualitative analysis of evidence.		
	i Evaluation uses credible forms of qualitative and quantitative data. It presents both output and outcome-level data as relevant to the evaluation framework. Triangulation is evident through the use of multiple data sources.	Yes	Although the analysis is detailed, it makes extensive use of all sources of data that were proposed in the methodology across many of the questions to be answered. For example, the first findings subsection (related to ICSE as a strategic enabler) combines secondary data on the background and evolution of ICSE as an organisational practice area with primary data on the perceived importance of ICSE both in other organisations today (via the inputs of the peer organisations that were included in the primary research) and within UNICEF via the feedback provided from the internal surveys, interviews and, where appropriate, the direct observations of the embedded evaluator (e.g. under question 2). This presents a well-reasoned and well-supported argument for the strategic importance of ICSE within UNICEF. This is a feature of many of the evaluation questions to be answered (although the later questions in the report are answered progressively more shallowly), also with useful presentation of individual case studies/examples to highlight useful or promising practices (e.g. Box 4).

	ii Findings are clearly supported by, and respond to, the evidence presented, including both positive and negative. Findings are based on clear performance indicators, standards, benchmarks, or other means of comparison as relevant for each question.	Partially	Overall, the evaluators make some good efforts to present analysis and evidence in the report from both primary and secondary sources, including data from interviews, the various surveys that were conducted and analysis of published documents and data (internal and external). This is a solid basis for the evaluation. However, a challenge faced in reading and understanding the report is that it does not present specific and clear findings per se. Each subsection starts with a short summary of the analysis - these are quite close to being findings, although they are not clearly noted as such, and do not quite constitute discrete and specific (as well as numbered) findings. The subsequent analysis is very narrative in nature, and readers must interpret this to determine the overall findings themselves. Firstly, this is not standard practice in evaluations - and would make any report challenging to parse and the essential direction of the evaluation difficult to grasp. Secondly, the nature of this evaluation, both subject matter and the design, is nuanced and particular efforts should have been made to ensure clarity. The evaluators do well to present the analysis against the evidence and the structure of the learning framework (as noted above), but how it is organised and presented could have been clearer to give the analysis more weight.
	iii Causal factors (contextual, organizational, managerial, etc.) leading to achievement or non-achievement of results are clearly identified. For theory-based evaluations, findings also analyse the logical chain (progression -or not- from implementation to results).	Yes	The exploration of causal factors is central to the nature of this evaluation. A considerable portion of the analysis is given over to presentation of evidence as to the underlying causes of ICSE challenges in UNICEF. For example, much of the analysis under the "enabling environment" dimension in the Learning Framework is given to determination of positive (enablers) and negative (barriers) factors that determine the overall effectiveness of ICSE. There is extensive analysis under each of these, with key drivers and barriers explained with reference to the evidence.
Question 14.	Does the evaluation assess and use the intervention's Results Based Management elements?		
	i Assessment of the adequacy of the intervention's monitoring system (including completeness and appropriateness of results/performance framework - including vertical and horizontal logic, M&E tools and their usage) to support decision-making.	Partially	Assessment of the "internal M&E mechanisms and tools" and the ICSE indicators in the integrated results and resources framework of the UNICEF strategic plan were the fourth and fifth (of five) objectives of the evaluation, and thus central to the evaluation. Although there was only a brief description of the relevant IRRF indicator in the background and context, there is a specific subsection within the analysis that covers the adequacy of the M&E system related to ICSE. While useful, this section is quite thin when considering that it purports to address two of the five evaluation objectives. Size-wise, it covers just over two pages, one of which is given to examples of external organisations, and evidence-wise, it relies heavily on what is presented in the IRRF (desk review) and the observations of the evaluator. While communication tools that can be leveraged for a monitoring function (such as uMessage, VivaEngage or ICON) are extensively discussed in other areas of the report (notably under Goal Area 3), more primary data on the perceptions of key stakeholders in UNICEF on the adequacy of monitoring and related needs might have been warranted. This said, the evaluators do make specific recommendations related to monitoring of ICSE, so the limited analysis is not necessarily reflected in the overall weight ascribed to this element by the authors.
SECTION F:	EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 10%)	63%	Comments on Rating
Question 15.	Do the conclusions clearly present an objective overall assessment of the intervention?		
	i Conclusions are clearly formulated and reflect the purpose and objectives of the evaluation. They are sufficiently forward looking (if a formative evaluation or if the implementation is expected to continue or have additional phase).	Yes	The conclusions section is strong, with clear arguments that reflect the overall purpose of the evaluation and the process of the developmental evaluation itself, which is less prescriptive than traditional evaluations. The conclusions are very much forward-looking insofar as the focus on where the gaps and opportunities lie for effective ICSE, as well as the key responsibilities for addressing the gaps or capitalising on the opportunities.
	ii Conclusions are derived appropriately from findings, and present a picture of the strengths and limitations of the intervention that adds insight and analysis beyond the findings.	Yes	In as much as the evaluation analysis and findings are weighted towards the strategic and management aspects of ICSE in UNICEF (i.e. under the "ICSE as Strategic Enabler" component of the Learning Framework), the conclusions are reflective of this - much of the discussion is on relatively high-level strengths and limitations against which the evidence is presented in the preceding sections. Therefore the conclusions are appropriately rooted in the findings and are well-crafted and insightful and should be of use to UNICEF.

Question 16.	Are logical and informative lessons learned identified? [N/A if lessons are not presented and not requested in ToR]		
	i Identified lessons stem logically from the findings, have wider applicability and relevance beyond the object of the evaluation.	Partially	Despite the section being entitled "Conclusions and Lessons Learned", there are no specific lessons presented in this section. It may be argued that the lessons are inherent to the conclusions, and indeed how the conclusions are written does reflect this to an extent, hence the criterion is rated as "partial". This said, there are a variety of useful lessons and good practices noted throughout the analysis/findings (as per the objectives of the evaluation) that could have been captured and codified in a more systematic manner here. Examples include: - On pg. 43, there is reference to a "UNICEF Internal Communication Playbook" that represents a standardized tool that could be of significant use to the wider organization. - The good practices on ICSE exhibited by the UNICEF Supply Division ("a bright spot") - The example of use of Vivo Engage in Gaza in 2023 (pg. 48) where poor follow-up on a request for information led to missed expectations and staff frustration.
	ii Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	No	Per the above, there is no specific lessons-learned subsection, and while the case may be made that the lessons are inherent to the conclusion statements, they do not meet the criterion of being "clearly and concisely presented" if so.
SECTION G: RECOMMENDATIONS (weight 15%)		63%	Comments on Rating
Question 17.	Are recommendations well grounded in the evaluation?		
	i Recommendations align with the evaluation purpose, are clearly formulated and logically derived from the findings and/or conclusions.	Partially	There are six recommendations made, and they are very well aligned to the evaluation purpose and are directly in line with the conclusions. The headline recommendation statements are clear, and are supported by strong summary arguments for their implementation that contextualise them well within the analysis. One element that, to this reviewer, seems to have been missed in the recommendations is the consideration of reducing the number of channels for communication and engagement and focusing on the highest-performers. Specifically, the data under Goal Area 3 (pp. 53-54) indicates that Viva Engage has the poorest engagement of all tools polled (15% of all staff, with a minimum of 30% deemed to be 'critical mass' for effectiveness) and that the various tools in use (ICON, uMessage, SharePoint) all fulfil similar roles and have higher engagement. Why then not recommend discontinuing of this unused channel? This is almost alluded to on pg. 55, which recommends "streamlining the number of tools or platforms available", but not made explicit and indeed engagement with the existing suite of platforms (including Viva Engage) is implicit within the recommendations.
	ii Recommendations are useful and actionable for primary intended users and uses (relevant to the intervention); guidance is given for implementation, as appropriate.	Yes	All of the recommendations make sense and are useful, and the priority that they are assigned is reflective of the analysis/findings and conclusions. Each recommendation has a number of suggested actions and opportunities that can help guide their implementation, as well as a brief note on the 'cost implications' for each, which is a most useful addition, given that cost considerations are a significant driver of management decisions (and indeed are noted in various places in the analysis). This is a very useful addition that could be mainstreamed in recommendations development for other evaluations.
	iii Process for developing the recommendations is described, and includes the Involvement of duty-bearers, as well as rights holders when feasible (or explanation given for why they were not involved).	No	Given the highly participatory nature of the evaluation, it is something of an omission that there is no description apparent of the process whereby the recommendations were developed. The evaluators were embedded with UNICEF for the evaluation process, at least one workshop was facilitated (for theory of change development), an evaluation reference group convened and key findings shared throughout via the learning memos. But there is no description of any participation or feedback processes for recommendations development, nor any explanation of why this might not have been the case.
Question 18.	Are recommendations clearly presented?		
	i Clear identification of groups or duty-bearers responsible for action for each recommendation (or clearly clustered group of recommendations). Clear prioritization and/or classification of recommendations to support use.	Yes	The presentation of the recommendations is very clear - there is a headline recommendation statement, with a short summary of the rationale for each, a priority (very high-high-medium), responsible units (within UNICEF) identified, a suggested timeframe (in months) and cost implications (general description of staff time and resources required - not a dollar amount). This is an optimal structure for the presentation (though some explanation of the priority levels would have been an enhancement).

SECTION H: REPORT STRUCTURE AND PRESENTATION (weight 5%)		86%	Comments on Rating
Question 19. Does the evaluation report include all relevant information?			
i	Opening pages include: Name of evaluated object, timeframe of the evaluation, date of report, location of evaluated object, names and/or organization(s) of the evaluator(s), name of organization commissioning the evaluation, table of contents -including, as relevant, tables, graphs, figures, annexes; list of acronyms/abbreviations, page numbers.	Yes	All of the required elements are present, with no omissions apparent. All page numbers line up, although the lists of figures/tables/boxes do not have page numbers or links assigned, which might be useful to navigate quickly to items of interest.
ii	Annexes include: terms of reference, evaluation matrix, list of interviewees, results chain/ToC/logical framework (unless included in report body), list of site visits, data collection instruments (such as survey or interview questionnaires), list of documentary evidence. Other appropriate annexes could include: additional details on methodology, information about the evaluator(s), etc.	Partially	The annexes include most of the required elements that are not otherwise in the main report. An omission are the survey tools. As noted above, the results of the surveys (with the questions before each figure) are included in Annex 5, but the tools themselves are not included, nor links to them online (if available). What is also not included in the annexes and might have been useful supplementary reading would have been the Learning Memos - various were produced as part of the evaluation and are referred to throughout the report, but not included in the annexes.
Question 20. Is the report logically structured?			
i	Structure is easy to identify and navigate (for instance, with numbered sections, clear titles and sub-titles, well formatted).	Yes	While the structure is generally straightforward, there is considerable narrative and many subheadings under the findings section - but subheadings are not numbered (only major headings) which does present a minor challenge to navigation, but is not egregious. The report is otherwise well formatted.
ii	Structure accords to UNICEF guidelines for evaluation reports: context, purpose and methodology would normally precede findings, which would normally be followed by conclusions, lessons learned and recommendations.	Yes	The structure is fully in accordance with the relevant guidelines.
Question 21. Is the report well presented?			
i	Report is of reasonable length; it does not exceed number of pages that may be specified in ToR.	Yes	The main report (excluding Executive Summary and Annexes) is 65 pages. There is no guidance specified in the TOR, but the length is quite reasonable - a good balance of detail and brevity.
ii	Report is easy to understand (written in accessible way for intended audience) and generally free from grammar, spelling and punctuation errors.	Partially	While overall the report is easy to understand, there are still a range of editing, spelling, grammatical and writing/style issues that should have been addressed pre-publication. Examples include: - Spelling mistakes - Repetitive language (see the first paragraph of section 1.1 - the same explanation of the ICON business unit is given in the same para). - Some of the language is jarring and/or colloquial or uses figures of speech or euphemisms rather than clear terms. - Use of jargon/buzzwords ("Silicon Valley", "Omnichannel approach", "sprints") without explanation. - Formatting of some of the footnotes is incorrect.
iii	Frequent use of visual aids (such as infographics, maps, tables, figures, photos) to convey key information. These are clearly presented, labelled, and referenced in text.	Yes	The report makes extensive use of charts, tables and boxes to present the quantitative data, text and examples/case studies. These add to the quality of the report, are easy to understand, and provide good visual impact in what is otherwise quite a narrative-heavy report. All are labelled correctly and displayed well.
SECTION I: EVALUATION PRINCIPLES (weight 10%)		47%	Comments on Rating
Question 22. 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?			
i	Reference and use of rights-based framework, and/or CRC, CCC, CEDAW and/or other rights related benchmarks in the design of the evaluation.	Not Rated	Given that this is an internal evaluation of UNICEF processes and systems, reference to these normative frameworks would not be expected, hence this criterion is not rated.
ii	Clear description of the level of participation of key rights holders and duty bearers in the conduct of the evaluation (for example, a reference group is established, stakeholders are involved as informants or in data gathering).	Yes	Given that this is an internal evaluation of UNICEF processes and systems, inclusion of and reference to rights-holders and external duty-bearer groups would not be expected. However, if one considers the involvement of UNICEF staff at different levels as the key stakeholders in this context, then there is extensive involvement of these at different levels of the organization as the evaluation was highly participatory, with the involvement of UNICEF staff in data gathering/key informants and with the inputs of the Evaluation Advisory Group.

	iii Language is empowering and inclusive, avoiding gender, heterosexual, age, cultural and religious bias, among others; use terminology of rights holders and duty bearers; data is disaggregated by marginalized group; differential results are assessed (distribution of results across different groups).	Yes	Given that this is an internal evaluation of UNICEF processes and systems, the requirement to reference these elements is minimal. Nonetheless, there was no evidence of discriminatory language in the evaluation. Responses to the all-staff survey were disaggregated by gender, but no other means of differential assessment was included. Disability is not mentioned in the report.
	iv Evaluation assesses the extent to which the implementation of the intervention addresses child rights and Leave No-one Behind (gender and other excluded and marginalized groups). It is disability inclusive.	Not Rated	Given that this is an internal evaluation of UNICEF processes and systems, reference to this would not be expected, hence this criterion is not rated.
Question 23.	Does the evaluation meet UN SWAP evaluation performance indicators? (Note: this question will be rated according to UN SWAP standards with detail provided below)	1	
	i GEEW is integrated in the Evaluation Scope of analysis, and evaluation criteria and questions are designed in a way that ensures GEEW-related data will be collected.	Partially integrated	There is a very light gender dimension to the evaluation, insofar as the risk to gender equality of poor ICSE is noted, and a consequence of it being gender discrimination and SEAH (the findings of the 2019 Independent Task Force (ITF) report on Workplace Gender Discrimination, Sexual Harassment, Harassment and Abuse of Authority were highlighted as (at least partially) driving this evaluation. However, the specific implications of ICSE for gender were not included in the scope of the evaluation, which is a gap, as some of the negative consequences of poor or limited ICSE (e.g. sexual harassment or discrimination, as well as gendered dimensions of power imbalances) are more pronounced for women and could usefully have been unpacked by the evaluators. Even an exploration of the extent of perceived importance of this as an issue amongst, say, management, would have been an interesting sub-finding.
	ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Not at all integrated	Following on from the light-touch reference to the gender implications of ICSE, there is no inclusion of these dimensions in the evaluation design, beyond generic references to normative frameworks and standards that the evaluation committed itself to.
	iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Not at all integrated	There is no discussion of gender or the gendered dimensions of ICSE in the analysis or conclusions/recommendations.

SWAP Rating Guidance

<p>i GEEW is integrated in the Evaluation Scope of analysis, and evaluation criteria and questions are designed in a way that ensures GEEW-related data will be collected.</p> <p>a. Does the evaluation assess whether sufficient information was collected during the implementation period on specific result indicators to measure progress on human rights and gender equality results?</p> <p>b. Does the evaluation include an objective specific to assessment of human rights and gender equality considerations or was it mainstreamed in other objectives?</p> <p>c. Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria?</p> <p>d. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation?</p>
<p>ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.</p> <p>a. Does the evaluation specify how gender issues are addressed in the methodology, including: how data collection and analysis methods integrate gender considerations and ensure data collected is disaggregated by sex?</p> <p>b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEWE considerations?</p> <p>c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility?</p> <p>d. Does the evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate?</p> <p>e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality?</p>
<p>iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.</p> <p>a. Does the evaluation have a background section that includes an intersectional analysis of the specific social groups affected by the issue or spell out the relevant normative instruments or policies related to human rights and gender equality?</p> <p>b. Do the findings include data analysis that explicitly and transparently triangulates the voices of different social role groups, and/or disaggregates quantitative data, where applicable?</p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described?</p> <p>d. Does the evaluation report provide specific recommendations addressing GEWE issues, and priorities for action to improve GEWE or the intervention or future initiatives in this area?</p>