

GEROS Evaluation Quality Assurance Tool

Version: September 2021

Evaluation of Integrated Approach to Programming in UNICEF Lebanon

REPORT RATING SUMMARY			
Overall Rating	56%	Fair	
●●●●●	Exceptional (96% - 100%)	5	
●●●●○	Highly Satisfactory (87.5% - 95.99%)	4	
●●●○●	Satisfactory (62.5% - 87.49%)	3	
●●○●●	Fair (35% - 62.49%)	2	Meets UNICEF/UNEG standards for evaluation reports in some regards, but not all. Decision makers may use the evaluation with caution. Substantive improvements in some areas are needed
●○○○○	Unsatisfactory (0% - 34.99%)	1	
REPORT DETAILS			
Title of the evaluation report	Evaluation of Integrated Approach to Programming in UNICEF Lebanon		
Report sequence number	Lebanon-59-2023-18800		
Region	MENA		
Year of report	2023		
Office	Evaluation Office		
Coverage (countries)	Lebanon		
ToRs present	Yes		
Date of review (dd/mmm/yyyy)	August 8, 2023		
Name of review firm	IOD PARC		
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)	Yes		
Evaluation object	Country Programme		
Evaluation type	Formative		
Evaluation strategy	Mixed methods		
Evaluation design (primary method used)	Non-experimental		
Evaluation level	Impact		
Geographic scope	National		
Primary SDG(s) covered (number)	1-8, 16		
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>			
Key strengths of the evaluation: >The report structure is clear and easy to follow. >Annex III clearly links the EQs to the evaluation criteria. >Data collection methods are appropriate. >Annex VI provides details of data collection tools and their related processes.			
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>			
Examples of improvements that could be made: >The scope of the evaluation requires more detail. The report should clearly indicate what components are included (i.e. if the evaluation covered only the six programmes listed in Figure 1 or if there are other LCO components covered). More details regarding the programmes covered would have been beneficial. >The evaluation report would have been strengthened by providing more details on linkages to relevant targets and indicators related to integration based on the integration programmes included under the evaluation. >The report could have more clearly explained why the evaluation was being undertaken 'now'. As written, it is unclear how the evaluation will feed into the already-adopted 2023-2025 CPD. >The report would have benefited from detailing how CQR was performed by an individual rather than the traditional method using multiple researchers to reach a consensus. >The findings section requires more detail from the KIIs, FGDs, and document review. Many of the findings appear to rely on only one piece of documentation or only one stakeholder's view instead of triangulating evidence. >The findings often lack sufficient levels of analysis and evidence to address the evaluation questions. Detailed feedback is given below. Overall, vague language mixed with a lack of clarity regarding where the evidence stemmed from leads to findings that require further strengthening. >The recommendations should be revisited to ensure that the responsible entity(ies) is/are listed for each recommendation. >The recommendations should include - at a minimum - a broad prioritization level (e.g., low, medium and high priority) for each recommendation. >The executive summary should provide more details on the findings so that conclusions and recommendations are more easily understood. >The report would benefit from more details around the 'established reference groups' as details are currently lacking.			
SECTION RATINGS			
SECTION A: EXECUTIVE SUMMARY (weight 5%)	67%	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).	Partially	The executive summary is about 4 pages in length. The executive summary cannot be used as a standalone document for the following reasons: The executive summary does not sufficiently summarize some of the criteria (e.g., coherence, effectiveness, efficiency) and would have benefited from offering more robust summaries. It is also difficult to understand how the ratings were determined. For example, the findings under the Relevance criterion paint a very positive picture. However, the rating listed under the Relevance criterion is 'partially satisfactory'. The executive summary also provides the conclusions and recommendations. However, there is some disconnect between these at times (e.g., conclusion 4 mentions M&E but the associated recommendation 4 does not address M&E) and there is lack of clarity for some recommendations (i.e., responsible entity(ies) for carrying out the recommendation). The two sentences under efficiency (conclusion 5) appear to contradict one another, and the information in this conclusion does not appear to be rooted in the information presented under the summarized efficiency findings.
	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	The executive summary includes all necessary elements. However, as mentioned above, the executive summary does not include enough detail to be a standalone document.
	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.	Partially	As mentioned above, the executive summary does not contain enough of a detailed summary in order to be considered a standalone document. The executive summary would have benefited from including more details on the findings so that conclusions and recommendations were more easily understood. The executive summary does not introduce new information from what is presented in the rest of the report.
SECTION B: BACKGROUND (weight 5%)			50%	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	Figure 1 clearly visualizes the covered interventions. The report lacks details on location(s), timelines, cost/budget and implementation status. The report states in a latter section that the evaluation covers 'all UNICEF Lebanon programme components, from 2018 to 2022, both at the national and local scales', but does not state whether the six programmes in figure 1 are the entirety of LCO programme components. More detail regarding the scope of the evaluation would have been beneficial to include.
	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).	Partially	The evaluation relates to the integrated approach to programming and not to a programme/intervention specifically, although it uses six integrated child-centered programmes to the relevance, effectiveness, coherence, efficiency, and sustainability of integration to the selected programmes of UNICEF, identify good practice and lessons learnt, and offer recommendations on how to strengthen integration across the length and breadth of UNICEF programmes in Lebanon. The report does not detail the rightsholders (beneficiaries) and duty bearers for each of these six programmes. Instead, the report offers a general outline of rightholders (beneficiaries) and briefly mentions partners in the section titled "UNICEF LCO's Approach to Integrated Programming". In this sense, the overall rightsholders (beneficiaries) are noted, but not per programme. Specific duty bearers are not listed. Geographic locations are not provided for the six programmes. There is no description of numbers reached by the six integrated child-centered programmes. At a minimum, the report would have benefited from clearly detailing each of the six programmes used to assess UNICEF's integrated approach in terms of their rightsholders (beneficiaries), duty bearers, and geographic coverage. The report would have benefited from more clearly stating what was included in the evaluation inclusive of a description of the rightsholders and duty bearers.
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.	Yes	The report clearly and concisely summarizes the country context, and how the context relates to the implementation of the intervention.
	ii	Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Partially	The report includes one sentence referencing the SDGs. The evaluation report would have been strengthened by providing more details on linkages to relevant targets and indicators related to integration based on the integration programmes included under the evaluation.
	iii	Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Yes	While more specific data could have been referenced, the country context provides sufficient details regarding the status of children in Lebanon.
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).	No	The report does not provide details related to the programmes visualized in Figure 1 that are included in the evaluation. More information on each of these six programmes including implementing partners, rightsholders, duty barriers and other stakeholders should have been included.
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	No	The report includes some information regarding the objectives of some interventions/programmes and lists some general activities, but does not detail the contribution of key stakeholders (including UNICEF) within the six programmes listed in Figure 1.
SECTION C:	EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)		67%	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.	Partially	The report is clear regarding the purpose/intended use and the key intended users. However, the question of 'why now' is not clearly articulated. While the report states: 'Based on identified lessons and best practices that should be continued/taken to scale or disseminated either within the programmes or more widely at the institutional level, UNICEF Lebanon seeks to improve programme integration in implementation of the country programme 2023 - 2025' it is unclear how the evaluation will feed into the country programming for 2023-2025 since the CPD has already been formulated. More details regarding how findings could influence the already-adopted country programme would be beneficial.
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).	Yes	The objectives of the evaluation are clearly stated in its own section. These are the same objectives listed in the TOR and inception report.
	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).	Partially	There is inconsistency regarding the scope of the evaluation and the author could have more clearly stated the scope of the evaluation. For example, there are some instances where six programmes are highlighted (e.g., Figure 1), instances where the terminology 'selected interventions' or 'selected programmes' is used, and other areas where the report states that 'all components' will be evaluated. This lack of clarity means that readers do not know what was included and what was excluded. Further, the objectives of the evaluation includes 'identifying any good practice within and outside integrated programming' -- this language is confusing as it appears that the scope of the evaluation is the six identified programmes from Figure 1 which included integrated programming approaches. So it is unclear how good practices outside of integrated programming could be highlighted as part of this evaluation.
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.	Not Rated	The objectives of the evaluation do not include testing the results chain.
	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).	Not Rated	The evaluation includes an intervention logic figure, but this does not include outputs and outcomes.
	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	The objectives of the evaluation do not include assessing a theory of change or results framework.
SECTION D:	EVALUATION DESIGN AND METHODOLOGY (weight 20%)		77%	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.	Yes	Annex III clearly links the EQs (although the column label still lists 'proposed EQs') to the relevant criteria and these are appropriate for meeting the objectives and purpose of the evaluation.

	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.	Yes	In addition to the EQs, the evaluation matrix lists sources of information and data collection methods.
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.	Yes	The evaluation design and set of methods (KIIs, FGDs, and an online survey) are adequate for this evaluation. However, the report does not clearly state which stakeholders took part in a KII or FGD. It would have been beneficial to include this level of detail.
	ii	Data sources are appropriate - these would normally include qualitative and quantitative sources (unless otherwise specified in the ToR) - and are all clearly described.	Yes	Data sources were appropriate and included survey responses, interview/FGD responses, and secondary documentation. Clear descriptions of the online survey, FGD and interview methods were included. The document and literature review was described within the 'data collection tools and processes' and 'data analysis' subsections instead of as a separate method with its own paragraph like the online survey, FGD and interviews. The amount of information on the documents and literature review is sufficient.
	iii	Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or, if not, provide reasons for this).	Yes	The sampling strategy is provided and includes a description of how diverse perspectives were captured: 1. The sampling strategy was based on ongoing discussions with UNICEF staff concerning the design and rollout of the approach and selected interventions. A purposive sampling approach was adopted to conduct the Focus Group Discussion and the Online Survey. Similarly, a purposive snowball sampling method was adopted for Key Informant Interviews. 2. The selection of respondents was done in consultation between the UNICEF team and the evaluator, taking into consideration the criteria set in the framework for the evaluation, such as a fair spread of meetings and interviews over the duration of the assignment, the profiles of persons and institutions met/interviewed ensuring a fair representation of the various target groups (UNICEF vs. implementing partners), as well as thematic/sectoral relevance as necessary."
	iv	Clear and complete description of the methods of analysis.	Partially	A complete and clear description of the data analysis methods was provided. However, the report would have benefited from additional details regarding how the CQR was performed by an individual rather than the traditional method using multiple researchers (i.e. CQR specifically uses multiple researchers to foster multiple perspectives and focuses on the process of reaching consensus through a systematic way of examining the results).
	v	Methodology allows for drawing causal connections between outputs and expected outcomes.	Partially	The methodology is designed in such a way that it would allow for drawing causal connections between outputs and outcomes. However, the findings often lack causal links.
	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 report highlights some of the limitations of data analysis and other limitations of the evaluation methods. The report contains a section related to constraints, adjustments, and mitigations. However, mitigation measures/adjustments related to bullet point 1 within the section are not offered.
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).	Yes	The report states that the ethical principles and standards defined by the UNEG were followed.
	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.	Yes	The ethical considerations for data collection were detailed and IRB approval was granted to the evaluation through HML.
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.

No	No innovative practices were utilized.
----	--

SECTION E: EVALUATION FINDINGS (weight 25%)		40%	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	Note that the EQ numbers in the findings section of the report do not match the EQ numbers in the evaluation matrix. Overall, there are many findings that lack sufficient levels of analysis to address the evaluation questions. Examples follow per the EQ numbers in the finding section of the report, not the EQs listed in the TOR, inception report, or evaluation matrix. EQ1. To what extent is UNICEF Lebanon's approach to integrated programming fit-for-purpose considering country context, programme's scope (types of programming across humanitarian and development) and implementation modalities? How relevant are different programme integration facets currently in use considering the overall humanitarian context/the very much vertical sectoral type of approach? EQ1 was supposed to investigate the appropriateness/relevance/fit-for-purpose nature of UNICEF Lebanon's approach to integrated programming given the country context, programmes' scopes, and implementation modalities. The findings under this section do not look at the approaches at the programme level (e.g., the six programmes referenced in Figure 1). Instead, the finding states that UNICEF staff were aware of the sectoral aspects of the CPD. The finding also states the expectations of sectoral programmes per the CPD but does not provide an evaluation of whether the approaches to integration within these sectoral programmes were appropriate/relevant/fit-for-purpose. The finding related to alignment with the country context/national priorities includes copy/paste views from key informants and lacks formulation of clear, triangulated findings. EQ2 findings includes one example from the ICWBP programme and it is
	ii Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Not Rated	There is no intervention ToC.
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.	Partially	The report uses qualitative and quantitative data. However, it is difficult to determine what data are feeding into findings as footnotes or references introducing quotations for example are often left out. Further, there are many findings that appear to rely on one document or one key informant. There are many cases where findings have not been triangulated.
	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	The findings are supported to the evidence presented, however - as stated in the cell above - it is difficult to determine what data are feeding into findings and there are many findings that appear to rely on one document or one key informant i.e., have not been triangulated. This evaluation did not have performance indicators/standards/benchmarks or other means of comparison.
	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).	Partially	Where appropriate, causal factors leading to achievements or shortcomings are briefly noted.
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.	No	While an evaluation subquestion under relevance was meant to look at the programmes' M&E systems ('To what extent are UNICEF programs internally coherent in terms of objectives, theories of change and M&E systems to facilitate integration?'), this was not assessed.
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	Conclusions are clearly formulated and reflect the purpose and objectives of the evaluation. Conclusions 2-6 are forward looking.

	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.	Partially	At times, it is difficult to understand which findings underpin the conclusions. It would have been beneficial to note which finding(s) informed each conclusion for clarity.
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	It is difficult to see how some lessons learned stem from the findings presented in the report (e.g., the lesson learned regarding data governance challenges). The lesson learned regarding building technical capacity of human resource is also disconnected from the findings as the findings section stated 'Staffing seemed to be "adequate and the skill mix seems right to meet the intended results". The 'evaluation of integrated programming' lesson learned is related to the evaluation specifically and not to integrated approaches. This should have been deleted. As presented, the lessons learnt are phrased in a way that allows for wider applicability.
	ii	Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Partially	The lessons learned listed contain vague language and it is difficult to relate them back to findings within the report. Sometimes the lessons appear to contradict findings (see comment above regarding capacity).
SECTION G:		RECOMMENDATIONS (weight 15%)	50%	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.	Yes	Recommendations align with the evaluation purpose and are derived from the findings and conclusions. Recommendation 1 'Clarify roles and responsibilities from the strategy, down to the programme and project levels' and recommendation 4 (To make ongoing efforts in program-level and project reporting effective, the intended approach to integrated programming should include a clarification of roles for it to be able to holistically deliver on the promise of the whole being more than the sum of its parts) appear to overlap or are duplicative.
	ii	Recommendations are useful and actionable for primary intended users and uses (relevant to the intervention); guidance is given for implementation, as appropriate.	Partially	The recommendations provide specific actions to be taken to fulfil each. The recommendations often provide the responsible party, but there are some that lack this level of detail. The report would have benefited from clearly stating the responsible party for each recommendation/recommended action.
	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	The report does not describe the process for developing the recommendations (except for stating that they are based on the findings). The report would have benefited from stating whether stakeholders were involved in the co-creation or validation of the recommendations.
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.	Partially	Some recommendations are clear regarding the responsible actor(s), while others do not include entity(ies) responsible (e.g., recommendations 1, 4, 6). There is no prioritization or classification for the recommendations, and this is stated in the introductory paragraph of the section: 'While many are categorized headings or descriptors, the recommendations reflected in Table 5 below are by no-means exhaustive, should not be seen as free-standing or mutually exclusive, and should not be read in rank order, as many should be carried out concurrently.' The report would have benefited from including a broad classification for recommendations, for example, 'low', 'medium' and 'high' priority levels.
SECTION H:		REPORT STRUCTURE AND PRESENTATION (weight 5%)	79%	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	The opening pages contain the information listed under this criterion.
	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.	Yes	The annexes include the TOR, TOC (intervention logic), evaluation matrix, list of documents reviewed, list of people interviewed, data collection tools and processes, the IRB approval letter, and a biography of the evaluator.

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).	Partially	The report is easy to navigate and the table of contents allows readers to click on the specific section for easy navigation. The report is somewhat well formatted but the inclusion of numbers for the sections and the addition of subsections - especially within the findings section - would have helped the reader more easily follow the structure of the final report. As submitted, the report is not well formatted in some areas (i.e., does not consistently use the same format to signify a new finding).
	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	Overall, the report follows a logical structure. The conclusions and recommendations are presented following the lessons learned and are in table format, but this does not impact the reader's understanding.
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 TOR states that the final report should not exceed 50 pages. The final report is 40 pages long.
	ii Report is easy to understand (written in accessible way for intended audience) and generally free from grammar, spelling and punctuation errors.	Partially	The report would have benefited from a thorough copyedit/proofreading as there are many sentences that contain grammatical errors. Small punctuation errors are also present, but these do not make the report more difficult to read. References to and labels of tables require updating.
	iii Frequent use of visual aids (such as infographics, maps, tables, figures, photos) to convey key information. These are clearly presented, labeled, and referenced in text.	Partially	The report appropriately utilizes visual aids. They are clearly labeled and are also referenced in the text. However, the specific table numbers for table references were not updated in the final version. For example: --the body of the report refers to Table 4 when the information being referenced is actually in Table 3. --the body of the report refers to Table 5 under "Conclusions and Recommendations", but the information being referenced is actually in Table 4.
SECTION I:	EVALUATION PRINCIPLES (weight 10%)	38%	Utilizing captions and cross-reference features of Word then updating labels 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.	No	The report should have included references to the CRC, CEDAW, CRPD, etc. Further, the report missed the opportunity to highlight the CCC as they relate to integration.
	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).	Partially	The evaluation methodology involved duty bearers but did not engage rightsholders. The report section titled 'quality control' mentioned 'established reference groups' but details of this group including its members and how many different reference groups were involved in the evaluation was not provided.
	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	The language used throughout the report avoids gender, heterosexual, age, cultural and religious bias. The report utilizes terminology such as rightsholders and duty bearers.
	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.	Partially	The report references the most vulnerable children in Lebanon throughout. However, the report did not assess the extent to which integrated programming addressed LNOB. The evaluation did not include any disability-inclusive language or questions.
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)	2	

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	<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? N/A</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? The evaluation states that human rights and gender equality considerations were mainstreamed into 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? The evaluation states that gender and/or human rights were mainstreamed into other evaluation criteria. However, only one question on the interview guide was included under relevance (To what extent did the approach adopt gender-sensitive, human rights-based and conflict-sensitive approaches?) and no findings related to gender-sensitive or human rights-based approaches were detailed in the findings section.</p> <p>d. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation? Yes, under relevance under relevance (To what extent did the approach adopt gender-sensitive, human rights-based and conflict-sensitive approaches?) but there are no findings related to gender-sensitive or human rights-based approaches in the findings section.</p>
ii	A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Partially integrated	<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? The report disaggregates key informants by sex. There is no other information regarding how data collection and analysis methods integrate gender considerations.</p> <p>b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEWE considerations? Yes, the evaluation employs a mix-methods approach.</p> <p>c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility? Many findings are derived from only one stakeholder. The report would benefit from including more details regarding where quotations are coming from and whether the views quoted are shared by more than one stakeholder.</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? The sampling frame included UNICEF and implementing partners. Gender balance was attempted but other diversity criteria were not considered.</p> <p>e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality?</p>
iii	The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Not at all integrated	<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? No.</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? No.</p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described? No.</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? No.</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.

- 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?
- b. Does the evaluation include an objective specific to assessment of human rights and gender equality considerations or was it mainstreamed in other objectives?
- c. Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria?
- d. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation?

ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.

- 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?
- b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEWE considerations?
- c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility?
- d. Does the evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate?
- e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality?

iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.

- 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?
- 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?
- c. Are unanticipated effects of the intervention on human rights and gender equality described?
- 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?