

# GEROS Evaluation Quality Assurance Tool

Version: September 2021

## Evaluation of the Child Protection System in Belize 2023

REPORT RATING SUMMARY			
<b>Overall Rating</b>	71%	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	
<b>Title of the evaluation report</b>	Evaluation of the Child Protection System in Belize 2023
<b>Report sequence number</b>	Belize/62/2023/19950
<b>Region</b>	LACR
<b>Year of report</b>	2024
<b>Office</b>	UNICEF Belize CO
<b>Coverage (countries)</b>	Belize
<b>ToRs present</b>	Yes
<b>Date of review (dd/mmm/yyyy)</b>	February 7, 2024
<b>Name of review firm</b>	IOD PARC

CLASSIFICATION OF EVALUATION REPORT	
<b>Management of evaluation (Managerial control and oversight of evaluation)</b>	Jointly managed with Country
<b>Unicef goal areas (Alignment with strategic plan priorities)</b>	
Every child survives and thrives	No
Every child learns	No
Every child is protected from violence and exploitation	Yes
Every child lives in a safe and clean environment	No
Every child has an equitable chance in life	Yes
Gender equality (cross-cutting)	Yes
Humanitarian action (cross-cutting)	No
<b>Evaluation object</b>	System
<b>Evaluation type</b>	Summative and formative
<b>Evaluation strategy</b>	Mixed methods
<b>Evaluation design (primary method used)</b>	Theory-based
<b>Evaluation level</b>	Output & Outcome
<b>Geographic scope</b>	National
<b>Primary SDG(s) covered (number)</b>	16

**EQA Summary:** *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.*

Overall this is a strong and well-written report, with much good analysis that will undoubtedly be of good use to UNICEF and national partners. Some key strengths and weaknesses are as follows:

- The executive summary is clear, concise, and follows the layout of the main report, serving as a useful standalone document summarizing main evaluation findings, conclusions, and recommendations, although it mirrors some of the lack of structured findings found in the main report.
- The background/context section offer a comprehensive overview of Belize's child protection context, including challenges faced and some quantitative data on key legal, policy, socio-economic, and institutional factors affecting child protection, although it lacks detailed resource allocation information and duty bearers are only partially identified, with limited discussion on their roles, responsibilities, and collaboration.
- The evaluation purpose and objectives are clearly defined, aligned with the evaluation framework, and identify key users of the evaluation.
- The evaluation design and methods are relevant and adequately described, although some areas lack detail, such as sampling strategies and data analysis.
- There is a clear reconstruction of the theory of change, well-linked to the evaluation framework and questions, enhancing methodological rigor.
- The evaluation questions are clear and realistic, based on DAC evaluation criteria, with modifications noted for clarity. The analysis mostly (though not fully) follows the questions as set out in the evaluation matrix in the annex.
- Findings are not articulated individually – the analysis section is quite narrative heavy, and the voices of rights holders are not present to the extent that the design would imply.
- Conclusions are strong and aligned with the evaluation criteria.
- There are extensive lessons-learned and good practices sections which offer valuable insights derived from the analysis.
- Recommendations are well-formulated, although not explicitly linked to conclusions, and provide clear implementation guidance for stakeholders.
- The report structure is sound, with all required elements and useful annexes, but minor formatting issues are present.
- Gender and equity considerations are quite well included, with specific efforts made in evaluation design, questions, and analysis, although limited data disaggregation is noted.

**Recommendations for Improvement:** *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.*

- Align the structure of the executive summary findings section with the evaluation questions for improved coherence and clarity.
- Provide more detailed information on resource allocations and budgets to contextualize findings within budgetary considerations.
- Improve the identification and description of duty bearers and stakeholders, specifying roles, responsibilities, and collaboration mechanisms.
- Clarify the geographical and chronological scope of the evaluation to ensure comprehensive coverage and understanding.
- Strengthen the description of methods and analysis, providing more detail on tools, work plans, testing, and data management.
- Incorporate some analysis of financial sustainability within the child protection system.
- Improve the analysis/findings by presenting standalone findings statements that reflect the considerable analysis that has been undertaken under each evaluation sub-question.
- Review the findings against the sub-questions to identify areas where analysis has not been presented
- Improve the integration of voices of rights-holders throughout the evaluation analysis to ensure diverse perspectives are adequately represented.
- Address any instances of recommendation-type language in the findings/analysis section, ensuring a clear separation between findings and recommendations.
- If available (and in line with the design commitments), enhance gender and equity considerations by incorporating more gender-disaggregated data and analysis throughout the evaluation, particularly in areas related to women and girls' experiences within the child protection system.

SECTION RATINGS			
<b>SECTION A:</b>	<b>EXECUTIVE SUMMARY (weight 5%)</b>	83%	Comments on Rating
<b>Question 1.</b>	Can the executive summary inform decision-making?		
	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 clear, concise (5 pages) and follows the layout of the main report. It can serve as a useful standalone document that summarises all of the main evaluation findings, conclusions and recommendations.

	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.	Partially	The structure of the executive summary is per the main report, although it is somewhat narrative-heavy (although it is less so than the main report, which has no itemised findings). As with the main report, it would have benefitted by being structured in accordance with the evaluation questions, with brief summaries of evidence under each - even simply summarising the conclusions (structured according to the evaluation questions) in the executive summary (instead of the findings) might be more coherent.
	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 summary largely achieves this, and is a good general summary that highlights the main achievements and challenges across the overall programme.
<b>SECTION B: BACKGROUND (weight 5%)</b>			71%	Comments on Rating
<b>Question 2.</b>		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	The evaluation is not an evaluation of an intervention per se - it assesses the entirety of the Belize child protection system, and is therefore not focused on a standalone or discrete initiative. It does, as described below, comprehensively describe the overall context of child protection in Belize, although no details on specific interventions or budgetary allocations (national or otherwise) are provided. Some detail on such resource flows would be useful context, particularly given that budgetary considerations form part of the evaluation framework/evaluation questions.
	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).	Yes	Sections 1 and 2 provide a comprehensive summary of the demographic, social, political, economic etc. background to Belize and its children, including the challenges faced. They include some quantitative data in terms of prevalence of some of the key issues faced by children (and as this is a national evaluation, these are the intended rights-holders to benefit from the CP system). As noted above, this is not an evaluation of a discrete initiative, so more granular breakdowns of these data are not warranted.
<b>Question 3.</b>		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 background/context section (section 2) provides a relevant, clear and comprehensive summary of the legal, policy, socio-economic, cultural, and institutional factors that affect child protection issues in Belize. For example, the text describes a number of legal frameworks governing CP (section 2.2). It also describes the socio-political/economic context, for example issues of poverty and wealth inequity, the diversity of the population, and the coordination mechanisms among various actors. The text also explains how the context relates to the implementation of the intervention, such as the gaps, opportunities, and recommendations to strengthen the child protection system.
	ii	Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Partially	The report only briefly discusses linkages with the SDGs in the background/context, with no mention of the specific SDG relevant to the evaluation area (Child protection, covered under SDG 16). Section 1, (pg. 2) notes that the Children's Agenda is aligned with the SDGs. The report analysis/findings do go into more detail, but not in the background section.
	iii	Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Yes	Section 2.2.1 and 2.2.2 discusses the background of rights and needs of children in Belize in considerable detail. The report identifies the main child protection issues in Belize, such as violence against children, child trafficking, child marriage and early unions (the difference between this and child marriage is not explained), cyber bullying, high prevalence of children with disabilities, and children in conflict with the law. It also provides data and evidence (in Section 2.2.2) drawn from the primary research to support the analysis of these issues, such as the prevalence of violent discipline, sexual abuse, and child poverty. The report also highlights the gaps and challenges in the child protection system, such as the lack of coordination, resources, standards, oversight, data, and participation mechanisms.
<b>Question 4.</b>		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	The report only partially identifies and describes duty bearers and other stakeholders, and the linkages between them, related to child protection in Belize. The report clearly highlights UNICEF and the Government of Belize as key actors, and briefly refers to some non-governmental organizations, civil society organizations, and faith-based organizations as service providers. They are noted as duty bearers in the theory of change description on pages 13/14 - and repeated word-for-word on page 15. The report does go into more detail of government agencies in Annex 9, but not specify the roles and responsibilities of other actors, nor how they collaborate and coordinate with each other.
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Partially	As above, the report briefly describes key stakeholders as part of the Theory of Change description, and of government stakeholders in Annex 9, but this is otherwise limited.
<b>SECTION C: EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)</b>			83%	Comments on Rating
<b>Question 5.</b>		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 report, in Section 3.0 clearly defines the evaluation purpose as to "assess the relevance, effectiveness, coherence and sustainability of the child protection system in Belize" and to "provide evidence-based recommendations for the Government of Belize and UNICEF to inform future programming and policy decisions". It grounds the object of the evaluation in the UNICEF Child Protection System Strengthening framework, and explains why the evaluation was needed and clearly identifies the key intended users of the evaluation, both duty-bearers and rightsholders, specifically the Government of Belize, UNICEF, civil society organizations, development partners and children and families.
<b>Question 6.</b>		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	The objectives and scope of the evaluation, as detailed on page 9, are clear, although quite ambitious (seven detailed objectives). The evaluation objectives and scope are aligned with the Terms of Reference (Annex 1), with some modifications made to aid clarity. However, this was not fully successful - the complexity of the objectives makes them difficult to understand. For example, objective three speaks about "guidelines", but is not clear whether the evaluation is supposed to provide/develop guidelines or whether there are pre-existing guidelines which UNICEF can build on. Some clearer articulation of the objectives could have been done at inception.

	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	The scope of the evaluation covers the following aspects: The scope, described in Section 3.2, is partially complete. The thematic aspect of the scope is clear, i.e. a focus on the child protection system in Belize, but the chronological and geographical aspects of the scope are not really discussed. It can be inferred from the background/context sections that the geographical scope is the entirety of the country, and, similarly, that there is no specific boundary on the chronological scope (the purpose being an assessment of the evolution of child protection in Belize to the time of data collection). However, this could be more clearly articulated. The scope section also provides detail on some of the limitations of the methodological scope, which is interesting, but should also be noted in the limitations section (discussed below).
<b>Question 7.</b>		<b>Is the theory of change, results chain or logic well articulated?</b>		
	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 reconstructed theory of change is well articulated (it is based on similar efforts in another evaluation in the region, which is a good practice). It provides a clear description of the intervention's intended results, as well as the assumptions, inputs, outputs, outcomes and impact that are relevant to the evaluation. It also shows the linkages between the intermediate outcomes and the higher-level outcomes of reduced violence, abuse, neglect and exploitation of children. The theory of change also identifies the key stakeholders, risks and enabling factors that may affect the implementation and effectiveness of the intervention.
	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).	Yes	The reconstructed TOC is presented in graphical format on page 12 and discussed in narrative across pages 11-15. There is some repetition of text in the narrative and poor formatting of some text (highlighted via in-line commentary), but it is otherwise good. An evaluation matrix has also been prepared based on the evaluation questions, presented in Annex 3.
	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.	Yes	As noted, the TOC was reconstructed as there was no existing TOC. This TOC is linked to the evaluation framework and questions and is cross-referenced extensively throughout the analysis and findings.
<b>SECTION D:</b>		<b>EVALUATION DESIGN AND METHODOLOGY (weight 20%)</b>	<b>64%</b>	<b>Comments on Rating</b>
<b>Question 8.</b>		<b>Does the evaluation use questions and the relevant evaluation criteria that are explicitly justified as appropriate for the purpose of the evaluation?</b> <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 evaluation questions for the evaluation are clear and realistic, with a good basis in the DAC evaluation criteria of relevance, effectiveness, coherence, and sustainability. The evaluation also covers the intermediate outcomes of the child protection system strengthening approach, such as legal and policy framework, governance and coordination structures, a continuum of services, etc. The report notes that the evaluation questions were modified during the inception phase and presents more detail in Annex 8. However, the main text does not provide any details on what those changes were or why they were made. A clear and complete description of the changes and their rationale would have enhanced the transparency and credibility 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.	Partially	Annex 3 includes the evaluation matrix/framework. It follows the evaluation questions and sub-questions faithfully (although the numbering of sub-subquestions in the matrix has errors) and has indicators, linkages to the TOC, data collection methods and analysis methods clearly noted. However, although it is largely robust, the breakdown from sub-questions to indicators in places does not fully capture all elements of the sub-questions. For example, evaluation sub-questions 4 & 5 (pg. 161) have a single indicator associated with them - related to sub-question 5 (the existence of CP safeguards) rather than 4 (integration of safeguards). Similarly for sub-question 10 (in the evaluation matrix), which is really two questions in one, but has a single indicator associated. While this is understandable in a complex evaluation, such simplification should have been done with respect to the EQs/sub-EQs themselves, rather than with the indicators. Should the evaluation commit to collect data only related to these indicators, the EQs/sub-EQs will be incompletely answered.
<b>Question 9.</b>		<b>Does the report specify adequate methods for data collection, analysis, and sampling?</b>		
	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 evaluation design and methods are relevant and adequate to meet the purpose of the evaluation, its objectives and are within scope, and are fully and clearly described. The evaluation design is based on a mixed methods approach that combines quantitative and qualitative data collection and analysis, using various tools such as key informant interviews, focus group discussions, online surveys (U-report), and desk review. This allows for triangulation of data sources and perspectives, as well as addressing the complexity and diversity of the child protection system and its stakeholders. Nonetheless, the methods are quite briefly described in the main text - there are essentially two short paragraphs that describe the methods and data sources, with another section on sampling. This is quite brief, particularly as it is not supported by an annex which might provide more detail (such as the nature of the tools - particularly U-report - the workplan, any testing/piloting exercises, survey platforms etc.).
	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	Despite the brief description of the methods in section 4, the data sources are quite appropriate to an evaluation of this nature, combining the perspectives of duty-bearers and rights-holders, primary and secondary, qualitative and quantitative methods. A wide range of tools was thus applied, which should ensure good triangulation between different sources and thus more robust findings.

	iii	Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or, if not, provide reasons for this).	Partially	The sampling strategy is provided in section 4.2.c. While adequate, it does not fully explain the criteria or rationale for the purposive sampling, nor does it provide any information on the representativeness or distribution of the respondents by district, sector, age group, or other relevant characteristics (the purposive sampling appears to be focused on KIs - the sampling strategy for FGDs and the surveys is less clear). The report also does not mention how the perspectives of children, especially those who are most vulnerable or marginalized, were captured or integrated into the evaluation. Therefore, the sampling strategy could be more transparent and comprehensive in describing how diverse perspectives are captured and how potential biases or gaps were addressed. Finally, the report states that "innovative approaches" were used to cover other stakeholders who could not be sampled for in-person or virtual interview, i.e. online surveys - however, this is not a particularly innovative method of data collection.
	iv	Clear and complete description of the methods of analysis.	Partially	The section on analysis (4.2.d) is quite brief, and while providing some detail on how data was managed and analysed (e.g. use of ATLAS.TI), disaggregation by location, gender and "type of respondent" (this latter is not explained further), and some details on linking data to evaluation questions and sub-domains of CPS outcomes. However, there is extraneous detail on the platform used for data collection (SurveyMonkey), some remarks about triangulation and noting the presence of desk review data (already explained above). Therefore this description is quite thin and incomplete - it lacks information on how data may have been managed, coded, cross-referenced etc. to ensure robust analysis.
	v	Methodology allows for drawing causal connections between outputs and expected outcomes.	Yes	Notwithstanding a lack of detail in the methods and analysis description, the use of multiple overlapping sources of primary/secondary; quant/qual sources of information, combined and cross-referenced as per the data analysis description, should lead to good capacity to draw causal connection between outputs and outcomes - as articulated in the results framework and the reconstructed TOC.
	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).	Yes	There is a clear description of the main limitations and challenges faced by the evaluators in section 4.3, including a table of the three primary limitations and mitigation strategies for each (somewhat repetitive with the narrative text). As noted above, the scope section also provides detail on some of the limitations of the methodological scope, which should have been noted in the limitations section, but is not.
<b>Question 10.</b>		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	Section 4.4 makes clear reference to the UNEG principles and notes the establishment of a reference group, conducting of an evaluability assessment and references "stakeholder consultation". No further specifics are noted in this very brief section, however. A more comprehensive and contextualised description should have been provided.
	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.	Partially	There is extensive description of the ethical safeguards applied in section 4.5. It describes a variety of measures applied to ensure ethical principles were adhered to. Examples include the obtaining of independent ethical clearance, the obtaining of informed consent from all participants and noting of "measures...in place to prevent children from being harmed through their participation, and redress mechanisms...instituted in collaboration with UNICEF and the Government of Belize" (Section 4.6.1). The research tools (Annex 2) include the consent statements and provisions applied before interviews or discussions. However, the report does not refer to the UNICEF procedures for Ethical Research Involving Children (specified in the TOR and germane as young people were involved in the research), nor does it explicitly address other issues relevant to methodology, such as respect for dignity and diversity, right to self-determination, fair representation, compliance with codes for vulnerable groups, or avoidance of harm.
<b>Question 11.</b>		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.	Partially	Although the evaluation claims to be innovative in the use of surveys for collecting data, such approaches are mainstream. The online feedback platform U-Report is somewhat innovative, as it takes advantage of social-media-type communication to solicit feedback directly from young people, but again, it is an established data collection tool and can only partially be considered innovative in this context.
<b>SECTION E: EVALUATION FINDINGS (weight 25%)</b>			<b>58%</b>	<b>Comments on Rating</b>
<b>Question 12.</b>		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	By and large, the evaluation does a good job in seeking to systematically follow the analysis plan per the evaluation matrix, with questions clearly stated in the preamble to the relevant subsections. However, the analysis is not summarised in standalone findings, rather there are blocks of narrative under each sub-section related to the sub-question, describing the landscape, highlighting strengths, weaknesses, achievements and challenges - but no overall findings that encapsulate all of these. This makes it difficult for the reader to establish what the tenor of the analyses are across each subsection. Annex 4 is a summary of findings with the key messages from the analysis presented under each evaluation question - the evaluators could consider copying these short summaries into the main report. Further, there are some gaps. For example, under EQ1 (Relevance), the evaluators omit sub-question #2 from the evaluation framework. Sub-questions 4 and 5 appear to have been combined and part of sub-question 10 is missed. From an evidence perspective, the evaluation also seeks to synthesise the available evidence well, with most evaluation questions answered with a combination of data. However, some areas are missed, for example that of the financial sustainability of the CPS. Although the scope of the evaluation explicitly notes covering "human, financial and infrastructure" components of the CPS (pg. 10, para 1), section 5.4.1.1 states that "The scope of this evaluation did not include assessing financial allocations or budgets quantitatively" - there is an obvious contradiction here which should be resolved. An evaluation of this nature should have included some analysis of funding flows, as part of both the background/context (noted above) and in the analysis under the specified and agreed question.
	ii Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Yes	The reconstructed TOC is clearly referenced throughout the analysis. Preambles to each analysis sub-section include a linkage of the question under discussion to the relevant TOC outcomes. This is a strong feature of the analysis, ensuring good methodological rigour.
<b>Question 13.</b>	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 relies heavily on qualitative and desk review data, more so than quantitative data. While triangulation is evident throughout the report, it is stronger under some questions than others. For example, the voices of respondents come through much more clearly under EQ3 (coherence), which provides more detail on the nature of the informants (e.g. "national child protection actor" vs "key informant") and introduces some direct quotations to highlight key points. Further, the voices of rights-holders are not present to the extent that they should. Notably under relevance and effectiveness (which, given the extensive questions in the FGD tools) provide good information on these aspects of the existing CPS with reference to the lived experience of these keystone stakeholders. Data from the U-report survey is referenced only once in the findings. These data could have been utilised more clearly, particularly using the direct voices (i.e. quotes) from rights-holders.
	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 evaluators do a good job of presenting evidence in relation to the analysis, with a combination of data/evidence from the desk review and perceptions of key informants and other stakeholders. Some of the analysis could be deeper - for example, under Q1 (relevance of institutional mandates), the report provides a summary of the mandates of key institutions related to CP, but does not discuss the relevance of these mandates, i.e. linking the description to established CP needs in Belize. All of the required data is present in the report, it simply lacks some evaluative analysis and judgement. Another example is under 5.1.2, which explores the relevance of institutional arrangements and structures. The "institutional arrangements and structures" are defined (for the purposes of the evaluation) in the evaluation matrix to include legal/policy, governance/coordination, case management, standards, SOPs, M&E etc - but the section discusses primarily the legal/policy framework.
	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	There is no discussion of results achieved with respect to CP in Belize. Thus, causal factors that may or may not lead to results cannot be determined. However, the identification of bottlenecks, hindrances and other factors related to is a central feature of the evaluation One small (and somewhat peripheral) issue is that some of the findings/analysis section contains elements that are phrased as recommendations, for example under 5.1.1 (pg. 28), where the evaluators speak about what "should" happen. The section should focus on findings that highlight deficits or issues to be addressed, and then be articulated as recommendations in the relevant section below.
<b>Question 14.</b>	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	M&E systems are assessed to some extent, but there is no overall TOC or results framework governing the entirety of the CPS in Belize, hence the evaluators cannot assess this aspect fully (noted in the limitations). Nonetheless, some aspects of M&E systems are discussed. For example, under Q1, the evaluators assess the relevance of the existing CPS indicators (noting some shortcomings) and under Q2 (5.2.2) the effectiveness of the "results-oriented approach", i.e. the M&E system (FAMcare), is explored, albeit briefly, with no analysis of any results framework or related data - the subsection relies on qualitative feedback related to strengths/deficits of the system only. While the evaluators acknowledge in the limitations/scope that such analysis could not take place, the evaluation would have been stronger if such analysis had been included.
<b>SECTION F:</b>	<b>EVALUATION CONCLUSIONS &amp; LESSONS LEARNED (weight 10%)</b>	<b>88%</b>	<b>Comments on Rating</b>
<b>Question 15.</b>	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).	Partially	Overall, the conclusions are quite strong, clear, reflect the evaluation purpose and objectives. They are structured in line with the evaluation criteria and it is clear from them what recommendations are in order. One issue with the conclusion related to sustainability (SQ1) is that it notes the difficult in the evaluation assessing financial aspects of the CPS - this should not really be a conclusion. It is more a limitation of the research. The conclusion could be more about the reason WHY financial commitments to CP is difficult (lack of a coherent approach, siloed working, poor sharing of financial data etc.) Further, there is a summary table of the evaluation criteria, questions and sub-questions in the preamble to the section - this is not necessary in the conclusions section and could be deleted.
	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	The narrative that accompanies the conclusions expands upon each with a good summary of the findings/analysis presented in the preceding sections, although as there are no individual 'findings' per se, they cannot explicitly link to individual findings, but do so to individual evaluation sub-questions, so are accurately derived. They serve as useful summaries of the analysis in the absence of the findings statements in section 4.
<b>Question 16.</b>		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.	Yes	There are quite extensive lessons-learned and good practices sections (6.2 and 6.3) which present some of the key learnings from the analysis. The lessons and practices that are presented are well derived from the analysis and findings in the main report and distill some key learnings that can be applied to both specific CPS implementation in Belize, but also to other countries and thus are relevant to UNICEF in a much wider context.
	ii	Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Yes	Although there is some repetition between the lessons and practices sections (e.g. around NGOs and the usefulness of the Hub), they are indeed clearly and concisely presented, with a headline statement supported by narrative analysis and examples. The formatting of both differs slightly (lessons are itemised "lesson 1", etc. whereas good practices are bulleted), but this is a minor issue.
<b>SECTION G: RECOMMENDATIONS (weight 15%)</b>			<b>75%</b>	<b>Comments on Rating</b>
<b>Question 17.</b>		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	The eight recommendations are generally well-aligned with the purpose of the evaluation, and the evaluators note their design in line with key rights principles. They are very clearly formulated, with the recommendations themselves including sub-recommendations and a number of supporting implementation guidance. However, they are not linked explicitly to the numbered conclusions or specific findings. In places in the findings/analysis section, some recommendation-type language emerges (e.g. for example under 5.1.1 (pg. 28), where the evaluators speak about what "should" happen.). Unless this is direct feedback from a primary source (e.g. a key stakeholder/s, strictly speaking this should be avoided in the main text and reserved for the recommendations section itself (the example cited is reflected in Recommendation 1).
	ii	Recommendations are useful and actionable for primary intended users and uses (relevant to the intervention); guidance is given for implementation, as appropriate.	Yes	The main recommendations and the sub-recommendations are clear and linked to specific actions that stakeholders can operationalise to ensure implementation. The implementation guidance provides additional detail and granularity that can assist stakeholders to apply them.
	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).	Partially	The process for developing the recommendations included a wide variety of institutional stakeholders across Belize. This is well explained in the preamble to the recommendations section. Interestingly, disagreement with some of the recommendations by some stakeholders is noted but could not be expanded on due to the onset of a natural disaster. Given that this no longer pertains, some revision of this aspect might be considered. Involvement of rightsholders in recommendations development was not explicitly noted or their absence explained.
<b>Question 18.</b>		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 recommendations are clearly presented in table format, with each recommendation addressed to relevant stakeholder and a primary responsible party assigned. A priority is tentatively assigned (as the evaluators have assigned UNICEF the responsibility for fully prioritising the recommendations, and a timeframe allocated (with a useful explanation of what each value means, e.g. how long is short-term/medium-term etc. - very positive).
<b>SECTION H: REPORT STRUCTURE AND PRESENTATION (weight 5%)</b>			<b>86%</b>	<b>Comments on Rating</b>
<b>Question 19.</b>		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	Yes	The structure of the report is sound, it has all required elements which are correctly formatted. The table of contents is present, and a table of figures/tables included (although the table of figures is incorrectly titled).
	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	There is a suite of useful annexes with the report, which includes the Evaluation Matrix, tools, stakeholder list, TOR, bibliography. The details and qualifications of the evaluation team could also be included, as could an expanded methodology annex detailing some of the absent elements noted above.
<b>Question 20.</b>		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	The structure is sound and easy to navigate. There are some formatting issues present that a final proof read could capture, e.g. subheading errors, spelling mistakes, formatting inconsistencies (some are marked via comment in the attached reviewed report), but they are minimal.

	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 report follows the UNICEF guidelines for structure.
<b>Question 21.</b>	Is the report well presented?		
	i Report is of reasonable length; it does not exceed number of pages that may be specified in ToR.	Partially	The TORs specify a main report of 50 pages, whereas the report itself is 79 pages (excluding executive summary). However, the introductory sections comprise 20 pages, with the conclusions, lessons and recommendations another 20 total - the findings/analysis is therefore approximately 30 pages, which is by no means excessive for the number of evaluation questions to be answered.
	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 is generally very well-written and easy to understand - it is of high quality. There are some formatting issues present that a final proof read could capture, e.g. subheading errors, spelling mistakes, formatting inconsistencies (some are marked via comment in the attached reviewed report), but they are minimal.
	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 good use of tables, figures and infographics. All are correctly labelled and referenced.
<b>SECTION I:</b>	<b>EVALUATION PRINCIPLES (weight 10%)</b>	<b>76%</b>	<b>Comments on Rating</b>
<b>Question 22.</b>	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.	Partially	The evaluation is clear in its intent to adopt a rights-based framework, articulated in section 4.2.d. This section emphasises relevant provisions of the CRC and the evaluation notes CEDAW and the SDGs elsewhere in the findings of the evaluation. Further, the evaluation questions and sub-questions are aligned with the child protection system components and sub-domains, which are based on human rights principles and standards. However, the evaluation design does not explicitly state how the rights-based frameworks and conventions will be operationalized or measured in the data collection and analysis processes beyond ensuring the voices of rights-holders (including people with disabilities) and disaggregation by sex. The evaluation design could have been more explicit in defining the methods that would capture the rights dimensions of the child protection system.
	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	Participation of rightsholders (both primary (children) and secondary (parents/caregivers)) and a duty-bearers is highlighted in the design of this evaluation, as discussed above. However, the challenges faced in obtaining a sufficient sample of rightsholders (noted in the limitations), may have had a significant limitation on the extent to which the report ensured that the inputs received were reflected in the analysis and findings. There are relatively few voices of rights-holders coming out in the analysis, which is disappointing, due to the importance of these in ground-truthing findings.
	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	Language is fully in accordance with expected standards and norms for this type of assignment. Data is not generally presented in a manner in which disaggregation is required, but inequities as they pertain to Gender Equality, Disability and Social Inclusion (GEDSI) are analysed and discussed across the analysis and also in a specific section (6.4), which is very positive.
	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.	Yes	As noted above, the report discusses specific vulnerabilities of groups throughout the analysis, and specifically within Section 6.4, which comprises analysis specifically related to GEDSI, and also presents specific recommendations for activities promoting GEDSI within the Belize CPS.
<b>Question 23.</b>	Does the evaluation meet UN SWAP evaluation performance indicators? (Note: this question will be rated according to UN SWAP standards with detail provided below)	7	
	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.	Satisfactorily integrated	Although there is no specific objectives related to gender and the objectives of the evaluation does not include mention of gender, the evaluation methods, scope and tools are reasonably-well oriented towards gender and equity considerations, with considerable efforts made in the evaluation design to note how the evaluation will be sensitive to these issues.
	ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Satisfactorily integrated	The approach notes specific gender-related aspects and this is also built into the evaluation questions, the evaluation matrix and data collection tools. Analysis is nominally responsive to gender, but in practice, there is limited scope for gender disaggregation of data as relatively little primary data related to the experiences of women and girls is presented.
	iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Fully integrated	While the main analysis is light on gender-specific elements, the standalone section on GEDSI (6.4), the conclusions themselves and the recommendations are well aligned towards including gender equity elements.

#### SWAP Rating Guidance

##### 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?