

GEROS Evaluation Quality Assurance Tool

Version: September 2021

Evaluation of the “Strategy for Exercising the Rights of the Child 2019-2023” in Montenegro

REPORT RATING SUMMARY			
Overall Rating	53%	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 the “Strategy for Exercising the Rights of the Child 2019-2023” in Montenegro		
Report sequence number	Republic of Montenegro 66-2024-19238		
Region	ECAR		
Year of report	2024		
Office	Montenegro Country Office		
Coverage (countries)	Montenegro		
ToRs present	Yes		
Date of review (dd/mmm/yyyy)	October 2, 2024		
Name of review firm	IOD PARC		
CLASSIFICATION OF EVALUATION REPORT			
Management of evaluation (Managerial control and oversight of evaluation)	Jointly managed with Country		
Unicef goal areas (Alignment with strategic plan priorities)			
Every child survives and thrives	Yes		
Every child learns	Yes		
Every child is protected from violence and exploitation	Yes		
Every child lives in a safe and clean environment	Yes		
Every child has an equitable chance in life	Yes		
Gender equality (cross-cutting)	Yes		
Humanitarian action (cross-cutting)	No		
Evaluation object	Strategy		
Evaluation type	Summative and formative		
Evaluation strategy	Mixed methods		
Evaluation design (primary method used)	Non-experimental		
Evaluation level	Output & Outcome		
Geographic scope	National		
Primary SDG(s) covered (number)	1, 4, 5, 6, 10, 16, 17		
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.			
<p>Overall, this evaluation report meets the minimum standard for quality evaluations. While there is some useful analysis which is reasonably well-grounded in the data and may be of use to the primary audiences (particularly if readers review the executive summary only), it has not fully delivered on the proposed design. A range of analytical elements are missing, and some mandated items, such as a lessons-learned section, have been entirely omitted, despite being a requirement in the TOR. Moreover, the report tends to be narrative-heavy, difficult to read due to the absence of findings statements, and some of the analysis areas are extremely thin. Specific strengths and weaknesses are as follows:</p> <p>Strengths:</p> <ul style="list-style-type: none"> •The report is structured in line with evaluation criteria with a logical flow between sections. •Detailed description of evaluation firm and team members' roles. •Useful suite of appendices that add additional information and context. •A solid analytical framework that links to evaluation criteria. •Detailed description of the evaluation's nature and key principles. •Clear articulation of the evaluation's purpose. •Good background and methods descriptions •Adequate standalone executive summary containing key findings and recommendations. <p>Weaknesses:</p> <ul style="list-style-type: none"> •English language inaccuracies and use of unexplained jargon. •Inaccurate terminology and poor formatting of diagrams. •The theory of change format is not well presented and its basis of development is not described. •Narrative heavy style, unsystematic reference to evaluation questions and no itemised findings. •Specific evaluation questions inconsistencies – not all sub questions are answered. •The analysis is superficial in places and does not present more critical insights to assist in future strategy formulation. •Absence of a lessons-learned section in the report. •Missing explanation of the process for recommendation development. •Some missing information, such as the report date and several unlisted acronyms. 			
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.			
<p>The major areas for improvement are as follows:</p> <ol style="list-style-type: none"> 1. A restructuring of the report to reflect the evaluation matrix more clearly, with ALL sub questions addressed 2. Findings itemised/findings statements included and clearly linked to the analysis 3. Expanded analysis of the final criteria (Sustainability and Impact) 4. Inclusion of a lessons-learned section 5. Thorough editing for language and style. 			
SECTION RATINGS			
SECTION A: EXECUTIVE SUMMARY (weight 5%)	83%	Comments on Rating	
Question 1. 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 summary is clear, concise, and precisely five pages long. It functions effectively as a standalone document, presenting the key findings, conclusions, and headline recommendations of the evaluation.	

	ii	Includes all necessary elements (overview of the intervention, evaluation purpose, objectives and intended audience, evaluation methodology, key conclusions on findings, lessons learned if requested, key recommendations) as per the ToR.	Partially	The structure of the executive summary is as per the main report, with a useful summary of the purpose/objective, background, methods etc. The findings in the summary are, essentially, a copy/paste of the conclusions, with the recommendations listed at the end. There are no lessons-learned (also missing from the 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.	Yes	Insofar as it the summary repeats the conclusions and recommendations found at the end of the report, i.e. as a standalone summary, it will suffice, although the major flaws of the main report are reflected in the summary also.
SECTION B: BACKGROUND (weight 5%)			71%	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	The description of the national strategy under evaluation is not overly complex, describes the purpose, objectives, the timeline, the implementing stakeholders and the budget. This is acceptable and not overly complex, which is a positive aspect. However, it does not provide much detail on the specific activities of the evaluation - what were the outputs, who implemented them, and when. Some more details on this, in brief, would have been useful. The language used in the description is somewhat unusual, with phrases like "appear to be" or "allegedly" introducing unnecessary ambiguity without clarification. Further, many of the statements in the background are not referenced, and where references are supplied, there is a mix of formats - in-line and footnoted. Many of the footnote references are simply "Data provided by UNICEF Montenegro" - this is not adequate for a reference. Clear and traceable references, formatted consistently, would have been useful.
	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	There is a sufficient description of the demographics of Montenegro in the background section, noting the age-breakdown of the population. The last pages of the background section (pages 14,15) include a description of some of the challenges faced by children in Montenegro across a range of dimensions. This background section specifically notes that the intended rights-holders of the strategy "appear to be all children, adolescents and young people in Montenegro", with parents/caregivers the indirect rights-holders. As this is a national strategy covering the entirety of the country, the description is accurate (despite the ambiguity of the phrasing, noted above). The different national agencies involved in children's rights are also described in terms of their participating in the strategy under evaluation. A comprehensive list is provided on page 11.
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	There is a solid description of the policy, political and other background to the strategy under evaluation. In particular, the evaluators describe three significant aspects to the policy background: 1. EU accession preparations which Montenegro is engaged in since 2010. 2. UN conventions and agreements. 3. National policies and laws governing the rights of children. All of these are described comprehensively, with some additional dimensions to the national context also provided.
	ii	Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Partially	The context section makes only one reference to the SDGs (pg. 11), but does not specify which goals and targets/indicators are relevant to the strategy. It notes that the findings section (5.2.1) explores this in more detail, but this does not appear to be the case. Addition of a more comprehensive explanation would have been welcome.
	iii	Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Yes	As noted above, there is a sufficient description of the demographics of Montenegro and the key issues related to children's wellbeing and development, although sources for the information are either lacking or could be improved.
Question 4.		Are key stakeholders, their relationships and contributions clearly identified?		
	i	Identification of implementing agency(ies), development partners, right holders, and additional duty bearers and other stakeholders; and of linkages between them (e.g., stakeholder map) (if relevant).	Partially	The background and context provides some explanation of the various duty bearers involved in children's development, but it is not particularly detailed. For example, on pg. 11, there is a list of 14 "thematic sectors" that are responsible for implementation of the strategy under evaluation (with one other national body and unspecified "overarching bodies" noted). This is not explored further, however. A more comprehensive stakeholder map, perhaps in a graphic form, with more detail in the annexes, would have been useful to illustrate the responsibilities for the strategy - including any linkages to civil society and/or international partners (UNICEF and others).
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Partially	As noted above, the description of stakeholders is cursory, with no detail on the specific contributions of these. The text notes that the responsibility for strategy implementation is "primarily the Government of Montenegro, line ministries and other national institutions". Financial detail is similarly generally specified, and indeed the evaluators note that more comprehensive financial details were "unavailable".
SECTION C: EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)			60%	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	There is a clearly articulated evaluation purpose, reflecting that in the TOR. As the strategy was at its conclusion and a new strategy under development, the need for the evaluation is clear. The purpose section also includes a list of the primary and secondary audiences for the evaluation, although some of the stakeholders are referred to in general terms and should have been made more explicit, i.e. organisation/entity names.
Question 6.		Are the objectives and scope of the evaluation clear and realistic?		

	i	Clear and complete description of what the evaluation seeks to achieve by the end of the process with reference to any changes made to the objectives included in the ToR (if applicable).	Partially	The purpose of the evaluation is relatively clear, articulated in section 2.1. Section 2.2 details the evaluation objectives, which are also reasonably clear, although the preamble text notes that the objectives are "in line with the evaluation TOR" - they are not quite the same objectives - some rearrangement and summarisation has taken place, but this is not made clear at the outset, which it should be. Most of this text is drawn from the inception report, so this is where the changes stem from. The IR is similarly unclear as to the changes.
	ii	Clear and relevant description of the scope of the evaluation: what will and will not be covered (thematically, chronologically, geographically with key terms defined), as well as, if applicable, the reasons for this scope (e.g., specifications by the ToRs, lack of access to particular geographic areas for political or safety reasons at the time of the evaluation, lack of data/evidence on particular elements of the intervention).	Yes	The scope of the evaluation is well-described, including temporal, geographical (it also discusses geographical sampling, which is not strictly required in this section but is acceptable) and programmatic/thematic. There are a number of limitations on the thematic scope (as articulated in the TOR) that are noted in the evaluation report, and were similarly highlighted in the Inception Report - this is acceptable, but a more careful and specific description of deviations from the TOR (i.e. a separate section) could have been included.
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.	Partially	As noted above, while the purpose and objectives of the strategy are clearly stated, the intended outcomes, results and activities are not explained in the relevant section. The evaluators do discuss the intervention logic of the strategy as part of the analysis and findings (section 5.1.2), and indeed a theory of change is reconstructed - but this is within the findings section, with no allusion or reference to it in the purpose/objectives/scope or the background.
	ii	Causal relationship between outputs and outcomes is presented in narrative and/or graphic form (e.g., results chain, logic model, theory of change, evaluation matrix).	Partially	As discussed above, a theory of change is reconstructed by the evaluators in the findings section only. While useful, it should be presented much earlier and the findings should focus on analysis of this only. Much of the relevant text in section 5.1.2, including the reconstructed ToC figure on page 34, could have been removed to the earlier sections.
	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 original strategy did not incorporate a theory of change, nor is there any reference in the TOR or evaluation report to a result framework (although outcomes, outputs and indicators are noted as existing). The evaluation is not theory-based (although the reconstructed theory of change is analysed/assessed in the findings), hence this criterion is not rated.
SECTION D:		EVALUATION DESIGN AND METHODOLOGY (weight 20%)	55%	Comments on Rating
Question 8.		Does the evaluation use questions and the relevant evaluation criteria that are explicitly justified as appropriate for the purpose of the evaluation? <i>UNICEF evaluation standards refer to the OECD/DAC criteria - Relevance; Coherence; Effectiveness; Efficiency; Sustainability; Impact (not all are necessarily relevant for all evaluations). Evaluations should also consider equity and leaving no-one behind, gender and human rights based approach (these can be mainstreamed into other criteria). Humanitarian evaluations should also consider Coverage; Connectedness; Coordination; Protection; Security.</i>		
	i	Evaluation questions and sub-questions are appropriate for meeting the objectives and purpose of the evaluation. The relevant criteria are specified and are aligned with the questions.	Partially	The evaluation questions have not been included in the main report, although they are referred to in the subheading for section 3.1. They have instead been moved to the annexes as part of the evaluation matrix. The remainder of section 3 discusses and specifies the evaluation criteria, which are present and suited to the evaluation. However, the previous section (2.3) on limitations discusses challenges with some of the criteria proposed in the TOR - section 3 would be a more appropriate place for this discussion, as well as any changes (or otherwise) to the evaluation questions. The absence of the evaluation questions in this section is somewhat puzzling as the section is very brief (<0.5 pages) and there are only ten questions - they could easily have been included without adding pages (if space-saving was the rationale for their omission). Finally, there is no discussion of any changes to the evaluation questions from those in the TOR - there are a variety of changes which were discussed and justified in the inception report, but not in this report - some text explaining the deviations from the TOR and the rationale for these should have been included.
	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	The evaluation matrix in the annexes breaks down the evaluation questions into two or more sub questions per question and associates a variety of indicators, data sources and data collection methods with each. A useful addition is a final column in the matrix on triangulation of data - if/how data from different sources will be compared.
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	Overall, the set of methods (desk/dataset review, key informant interviews, rights-holder survey) are adequate for this type of evaluation, if applied to a sufficiently robust sample of stakeholders. The methods are reasonably well, if briefly, described, with the essential details included in the main text, and links to the tools in the annexes.
	ii	Data sources are appropriate - these would normally include qualitative and quantitative sources (unless otherwise specified in the ToR) - and are all clearly described.	Partially	The sources of data are (a) institutional stakeholders (i.e. duty-bearers) and (b) rights holders. These are appropriate for an evaluation of this nature. However, they are incompletely described - particularly the duty-bearers. While the methods section refers to a stakeholder mapping exercise, this map is not provided, and institutional stakeholders are presented via a generic description (e.g. NGO partners are not itemised or named).

	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 only partially described. As noted above, the stakeholder mapping outputs are not included (even as an annex), so it is not possible to see what the sample universe looks like. Nor is the target sample shared - the report simply notes the ultimate number of individuals interviewed (24 national, 29 local/municipal - no sex disaggregation) and rights-holders (79 children - no sex or age disaggregation). Without knowing the overall sample universe, it is not possible to determine how representative or robust this sample is (although it appears to be on the small side given that this is an evaluation of a national-level strategy with at least 10-15 separate institutions involved, on a rough count). The report does note criteria for selection of municipalities in the sample, which is positive, and also that children representing "vulnerable groups" were solicited, but how this was done is not explained.
	iv	Clear and complete description of the methods of analysis.	No	There is a very brief description of the analysis - essentially the report states that "identifying themes, coding and clustering data" (pg. 25) is how data was analysed (this is a repetition of a similar sentence on pg. 23, para 5). This is not adequate. A more comprehensive description of this process is required.
	v	Methodology allows for drawing causal connections between outputs and expected outcomes.	Partially	The evaluation matrix presents a reasonable breakdown of the evaluation questions into logical components via the sub questions and indicators, with data sources/methods allocated to each. This is standard evaluation practice in ensuring logical links and establishing causality. Certainly some of the interview questions address the issue of causality (questions under effectiveness and impact), but the data tools presented in the annex are not cross-referenced with evaluation questions or sub questions, so it is not clear if/how the data tools are linked to the evaluation matrix and if they are adequate to adequately determine causal connections.
	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	Section 2.3 discusses a range of limitations/constraints faced by the evaluators. These were mainly in the area of sufficiency/availability of data, plans or strategies that impacted on the ability of the evaluators to determine key aspects of the evaluation (such as the absence of monitoring data impacting on assessing effectiveness). The section contains some useful information, but is presented in a narrative format - it would be more useful to present it in a table format, with risks in one column, mitigation strategies in another, and the ultimate realisation of risks/impact on the evaluation in a third.
Question 10.		Are ethical issues and considerations described? The evaluation should be guided by the UNEG ethical standards for evaluation. As such, the evaluation report should include:		
	i	Explicit and contextualized reference to the UNEG obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability) and/or the principles in the 2020 revised UNEG Ethical Guidelines (integrity, accountability, respect, beneficence).	Partially	The ethical considerations section (2.4) clearly notes the UNEG obligations of the evaluators, to the extent of excerpting several paragraphs from the UNEG Ethical Guidelines for Evaluation (these are not really necessary and could be deleted). However, this text refers to the 2017 edition of the guidelines - this should be updated to the 2020 version of the text (or just deleted). Notwithstanding this, the section explicitly notes the four principles and correctly cites the 2020 guidelines as the source for these. Further, the section notes and contextualises the obligations of impartiality and independence (among others), with those of credibility and accountability not described (conflict of interest is noted under impartiality).
	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	The ethics section describes well how a number of principles related to ethical evaluation are contextualised in the research. The evaluation also benefitted from the use of an Institutional Review Board for ethical approval. One issue that is not well explained (or potentially understood - although it may be a language issue with the authors) is that of consent of children, discussed in the ethics section on page 20. The evaluators assert that children could "provide consent for themselves". By definition, children cannot "consent". In such cases, children may give "assent" and formal "consent" obtained from designated caregivers. The evaluators should take particular care with terminology and language here, given the importance of ethical standards when researching among children.
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 applied in the evaluation.
SECTION E: EVALUATION FINDINGS (weight 25%)			50%	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.	No	The evaluation does not adopt a structured approach to presentation of findings in accordance with the evaluation matrix. While the findings section is organised by evaluation criterion, the evaluation questions and sub questions are not included, and there are no specific findings statements. Further, many of the areas of analysis proposed in the evaluation matrix are missing. For example, under Coherence, the analysis covers internal and external coherence, and predominantly discusses the coordination efforts of the strategy implementation (related to sub question 3.1 and EQ 4). However, the evaluation matrix includes a number of key dimensions that are not covered - inclusion of marginalised groups, gender sensitivity and child participation (this last is noted very briefly) in line with sub questions 3.2, 3.3, 3.4. Similarly for Effectiveness - the report does a good job of exploring the attainment of the strategy objectives (sub question 5.1), but then largely skips the remaining sub questions for the rest of this section. Similarly for Efficiency - the section discusses well the adequacy of the strategy budget/budget planning (sub questions 7.1, 8.1) but miss 7.2 and 8.2. For Impact, there is very little analysis (<1 page, all related to sub question 9.1), and while the evaluators note that measurement of impact is a limitation of the evaluation, the related sub question (9.2) should have been possible to answer as it is related to stakeholder perceptions of impact - such perceptions are easy to solicit in data collection. For Sustainability, both sub questions are answered, but very briefly (approx. 0.5 pages).
	ii Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Partially	As discussed above, a theory of change is reconstructed by the evaluators in the findings section. Whereas it should also have been presented in the background/context sections earlier and the findings focus on analysis only. Notwithstanding this aspect, the evaluators do present and discuss the reconstructed TOC in 5.1.2, although their analysis is more focused on the rationale for why a TOC was not prepared, which somewhat misses the point of a reconstructed TOC in an evaluation.
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	In general, the evaluation tries to use various forms of evidence - primary from interviewed stakeholders, secondary from indicators, reports etc., leading to reasonable levels of triangulation where the analysis is comprehensive. The analysis under the themes is presented in accordance with this evidence in most instances. A positive (if overly-detailed) example of the use of clear indicators is under Coherence, where a range of strategies/policies from the Montenegro Government are analysed for alignment with the strategy under evaluation. Another example can be seen under Effectiveness (5.3.1), where achievement of the strategy objectives is analysed with reference to a range of primary and secondary data. This said, the absence of findings statements means that there is no clear summary of what the analyses are saying, which makes the evaluation challenging to read.
	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	As discussed in detail above, the evaluation does not have clearly presented and itemised findings - there is, instead, a stream of analysis and discussion, which, while relatively well-rooted in evidence in places, there is little in the way of systematic or clear performance benchmarks, indicators or other comparative means. For example, the reconstructed TOC is not used to provide a basis for analysis of strategy achievement or performance, despite a variety of outcome measures being included therein. For some of the evaluation areas, the analysis is extremely thin, e.g. Impact and Sustainability, both of which are less than one page each, very much not in line with the plan outlined in the evaluation matrix. Finally, there is extensive use of ambiguous qualifiers (the phrases "appears to" and "seems to" are used extensively when presenting analysis). While this may be a language issue, it gives the impression that the evaluators are not sure of their analysis or evidence and should be avoided wherever possible.
	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	The identification of causal factors is an explicit part of the analysis under Effectiveness (EQ6, sub questions 6.1-6.3) and is presented in a general way over two pages in section 5.3.2. This is useful analysis. There are also some positive examples of causal analysis outside of this section. For example, under Coherence, the evaluators seek to explore the factors that have led to limited cooperation/coordination, which is positive. However, there are other areas where important causal analysis has simply been skipped - for example, under Effectiveness, a detailed analysis of the achievements/progress in regards to child rights in Montenegro is undercut by the inability/unwillingness of the evaluators to explore causal factors, simply stating "Whether the above can be attributed to the child rights Strategy...is however hard to establish." This is not a sufficient explanation.
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.	Yes	Insofar as the strategy under evaluation did not have a monitoring system in place, and that this absence is noted widely through the report as a limiting factor not just for the quality of the strategy implementation, but for the evaluation itself, it is well covered.
SECTION F: EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 10%)		25%	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).	Partially	The conclusions and recommendations section presents conclusions under two broad subheadings which purport to cover all findings - essentially strategy design and strategy effectiveness. They do not cover issue of coherence or sustainability. This said, the conclusions are reasonably clearly written and do reflect the core evaluation objectives. However, the conclusions are not itemised as individual sentences (or bullets, or even a headline conclusion statement for each), but as a single narrative block of text under each criterion - this is quite difficult for the reader to parse and is not particularly clear for reading. They are accompanied by the recommendations of the evaluation in one section. While this is not a typical approach to formulation of conclusions, it seems to work for the report, insofar as they do reflect the purpose and objectives of the evaluation. Some more structure and comprehensiveness to the overall conclusion - as with the findings (with more attention paid to the EQs) - would have been more useful, however.
	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	As presented, the two narrative-style conclusions do reflect the analysis in the main findings section, with plenty of critical feedback presented in this narrative, reflecting the shortcomings noted in the main report (e.g. lack of leadership, monitoring data, coordination issues etc). The overall impression (as can be said for much of the report itself) is that the conclusions are somewhat superficial and do not present more critical insights that could be of use to UNICEF and the Government of Montenegro to assist in future strategy formulation. Further, they repeat one of the major issues of the analysis, in that they miss much analysis that was proposed in the evaluation design and do not without reference the specific evaluation questions/sub-questions from which they are drawn. As can be said for the findings/analysis, the report conclusions would have benefited from a more systematic, comprehensive targeted and focused design.
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.	No	There is no lessons-learned section in the report, despite this being clearly noted as a requirement in the TOR, and the Inception Report reiterating this commitment (the IR notes that the Lessons Learned section will precede the Conclusions and Recommendations section).
	ii	Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	No	As above, no lessons learned section is included.
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.	Partially	The recommendations are listed immediately after the relevant conclusions, which is not a typical means of presentation, but adequately represents them and ensures that the reader can clearly understand the logic behind them (i.e. the conclusions from which they are derived). Insofar as the evaluation report is structured (i.e. only partially in line with the evaluation matrix), the recommendations mostly align with the findings and conclusions, with some exceptions. For example, the first conclusion statement notes the absence of a fully formulated theory of change (although the findings state there was no theory of change, which is not quite the same thing), but there is no associated recommendation to create or operationalise one (or at least use the reconstructed TOC prepared by the evaluators in the findings).
	ii	Recommendations are useful and actionable for primary intended users and uses (relevant to the intervention); guidance is given for implementation, as appropriate.	Partially	Many of the recommendations have merit and are useful in a general sense and, while actionable, do not have specific action points or sub-recommendations targeted at specific stakeholders. There are elements of this within the recommendations - specific stakeholder institutions are named, and examples of actions are provided (e.g. the design and operationalisation of action plans under R3, the use of gender indicators under R4), but there is a mix of conclusion-type language mixed with the recommendations - again, the very narrative style of the report works against it in this regard.
	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	There is no discussion about the process for recommendation development.
Question 18.		Are recommendations clearly presented?		
	i	Clear identification of groups or duty-bearers responsible for action for each recommendation (or clearly clustered group of recommendations). Clear prioritization and/or classification of recommendations to support use.	Yes	The recommendations do note the specific duty-bearer stakeholders to which they are targeted, although the narrative approach makes parsing specific actions to specific stakeholders difficult. They are also prioritised into high-medium-low, which is useful.
SECTION H: REPORT STRUCTURE AND PRESENTATION (weight 5%)			64%	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	Most of the required information is present per the criterion, with the exception of the date of the report and several acronyms are missing from the list on page 4 - some of the instances of these are highlighted in the report via in-line commentary.
	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 full suite of annexes provided which are useful and add value to the report.
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	Overall, the report is somewhat accessible, but numbered sections are not in line with evaluation questions or sub questions, making cross-checking challenging. There are a range of editing issues with the report. Many are captured via in-line commentary, with the main examples being: - Considerable English language inaccuracies - Unnecessary use of subheading in section 3 and 4 - there is only a single subsection in each, so the numbered subheadings are redundant. - Some disjointedness in sections, notably the methodology (section 4). In this, the data collection instrument description should come first, then the sampling, and finally the analysis description. - Acronyms/abbreviations missed, or not spelled out in first usage - Incorrect spelling of numbers (0-9 should be spelled out, 10+ as numerals) - Use of jargon or technical terms (PISA, Venice Commission) without explanation
	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.	Partially	The report accords with the UNICEF guidelines as laid out in the TOR and IR, with the exception of the missing lessons-learned section. The TOR also specifies that a Theory of Change will be included with the recommendations section - this is present instead in the findings section. Revisions to the report should include reorganisation to meet the requirements of the TOR.
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 requirement specified in the evaluation ToR is that the report should be of the minimum length of 50 pages (plus annexes). At 54 pages, this criterion has been adhered to. Notwithstanding the absence of specific areas of analysis, the length is not excessive.
	ii Report is easy to understand (written in accessible way for intended audience) and generally free from grammar, spelling and punctuation errors.	Partially	As noted above, the writing is generally comprehensible, but there are a considerable number of unusual or inaccurate terms or language errors. Comprehensive review from an English-language editor to catch and revise these is required.
	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.	No	The report has one table (in the annexes - not really a table) and one graphic (the reconstructed TOC). There is a noted absence of other visual aids, which adds to the issue with the report being narrative-heavy and difficult to read. In particular, there are no results from the survey presented in visual form (as would be expected from a survey).
SECTION I:	EVALUATION PRINCIPLES (weight 10%)	60%	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.	Yes	There design of the evaluation makes particular reference to the CRC which is central to the strategy of Montenegro in child rights. Assessing the implementation of the CRC provisions also forms part of the evaluation analysis.
	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	Rights holders (children) were involved in the data collection component of the evaluation - albeit only via a survey administered via implementing partner NGOs. Given that the initial strategy design involved consultations with children, more substantive involvement of children in the evaluation would have added significant value - even (if possible) involvement of the original cohort of children that contributed to the design of the strategy to assess their perspectives on that and any subsequent involvement or implementation would have been very useful.
	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).	Partially	Insofar as the analysis is presented (a narrative focus, no real quantitative analysis that could be disaggregated), the language is inclusive, with no indication of bias. However, the language of rights-holders and duty bearers is not used beyond the description of the strategy in section 1.2.
	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 evaluation does a good job of including consideration in the analysis of marginalised groups (child rights are central to the strategy, so this is necessarily fully integrated). There is consideration of children with disabilities, some discussion of gender dimensions, and other marginalised groups (notably Roma and Egyptians) are included in the analysis, though only gender considerations are included in the recommendations, which seems to be an omission given the ongoing challenges that these other groups face.
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)	5	
	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	Gender equality has been mainstreamed into the evaluation design, with the relevant guidance on this from UNICEF cited in the introduction, forming an element of one of the evaluation objectives and some key gender markers described in the context.
	ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Partially integrated	Despite the mentioned commitments to gender equity/equality cited in the design preamble text, the methods and tools do not mention gender equity or gender equality. The evaluation matrix does include specific sub questions and indicators that purport to explore gender dimensions of the strategy (e.g. sub questions 2.1 and 3.3) and the interview tools include several (i.e. 2-3) questions with a specific gender dimension, but it is not mainstreamed comprehensively.
	iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Satisfactorily integrated	The findings do include some analysis in relation to gender, derived from the data collected, and gender is specifically included in two recommendations (R3, R4).

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?