GEROS Evaluation Quality Assurance Tool

Version: February, 2020

UNFPA-UNICEF Global Programme to Accelerate Action to End Child Marriage

REPORT RATING SUMMARY		
Overall Rating	**94%**	**Highly Satisfactory**

● ● ● ●	Highly Satisfactory (87.5% - 100%)	4	Exceeds UNICEF/UNEG standards for evaluation reports and decision makers may use the evaluation with a high degree of confidence
● ● ● ─	Satisfactory (62.5% - 87.49%)	3	
● ● ─ ─	Fair (34.76% - 62.49%)	2	
● ─ ─ ─	Unsatisfactory (0% - 34.75%)	1	

REPORT DETAILS	
Title of the evaluation report	UNFPA-UNICEF Global Programme to Accelerate Action to End Child Marriage
Report sequence number	Evaluation Office/65/2019/16498
Region	EO
Year of report	2019
Office	Evaluation Office
Coverage (countries)	Global
ToRs present	Yes
Date of review (dd/mmm/yyyy)	15-01-2020
Name of review firm	AIMS

CLASSIFICATION OF EVALUATION REPORT	
Management of evaluation (*Managerial control and oversight of evaluation decisions*)	Jointly managed with one or more UN agencies
Unicef Goal Areas	
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	No
Gender equality (cross-cutting)	Yes
Humanitarian action (cross-cutting)	Yes
Evaluation object	Thematic area
Evaluation type	Summative and formative
Evaluation strategy	Mixed methods
Evaluation design (primary method used)	Theory-based
Evaluation level	Output & Outcome
Geographic Scope	Multi-region/Global
Primary SDG(s) covered	5.3

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.*

This is a complex evaluation of a complex project that is completing its initial phase. Ending child marriage involves some issues that are simple to measure, like adoption of legislation, and issues that are less easy to measure, like changes in community-level conditions due to poverty. Withal, the evaluation does a good job of mobilizing data to determine the extent to which Phase I has been successful in achieving its expected outcomes. The evaluation uses a carefully-developed theory of change and notes that this has been less used at the country level. The methodology relies heavily on document analysis, but supplements this with a large number of interviews. There were three field visits and in them focus groups with beneficiaries at the village level. How the specific villages to visit were chosen is not described, however, although it seems that they were suggested by country teams. The conclusions flow from the very detailed findings and the recommendations are mostly for the next phase and flow from issues that were observed. Gender issues were adequately addressed. The Executive Summary was longer than is generally recommended.

Recommendations for Improvement: *The rater will identify topline recommendations that would improve the evaluation, and be specific to the sections of the report where shortcomings were found. Cite resources that may assist evaluation managers in overseeing future evaluations as relevant.*

One weakness of the evaluation was a lack of information about how villages were chosen for visits and focus groups. This could have been better described because one of the problems noted by the evaluation is connecting child marriage reduction with broader issues of gender at the community level and having a clear sense of how the data were gathered would improve the credibility of the evaluation. It is also suggested that the Executive Summary be limited to approximately 5 pages as succinct summaries are more likely to be read in full (the maximum length should be specified in the ToRs).

SECTION RATINGS			
SECTION A:	**EXECUTIVE SUMMARY (weight 5%)**	**83%**	Comments on Rating (include explanations for any criterion not rated)
Question 1.	Can the executive summary inform decision-making?	83%	
	i An executive summary is provided that is of relevant conciseness and depth for primary intended users (Maximum of approximately 5 pages unless otherwise specified in ToR)	Partially	All of the issues are covered, but the Summary is 8 pages (nine in French and Spanish) and could be made more concise.
	ii Includes all necessary elements (overview of the intervention, evaluation purpose, objectives and intended audience, evaluation methodology, key findings, key conclusions, key recommendations) as per the ToR.	Yes	All of the elements are found in the Summary.

	iii	Includes all the necessary information to understand the intervention and the evaluation AND does not contain information not already included in the rest of the report	Yes	The executive summary provides sufficient information to understand both the programme and the evaluation approach and findings. In addition, there are versions in English, Spanish and French.
SECTION B:		**BACKGROUND (weight 5%)**	100%	Comments on Rating (include explanations for any criterion not rated)
Question 2.		Is the object of the evaluation clearly described?	100%	
	i	Clear and relevant description of the intervention, including: location(s), timelines, cost/budget, and implementation status	Yes	The project is well-described.
	ii	Clear and relevant description of intended beneficiaries by type (i.e., institutions/organizations; communities; individuals…), by geographic location(s) (i.e., urban, rural, particular neighbourhoods, town/cites, sub-regions…) and in terms of numbers reached (as appropriate to the purpose of the evaluation)	Yes	The intended beneficiaries (young women, their households and government and NGO programs dealing with child marriage) are described.
Question 3.		Is the context of the intervention clearly described?	100%	
	i	Clear and relevant description of the context of the intervention (relevant policy, socio-economic, political, institutional, international factors) and how context relates to the implementation of the intervention	Yes	This is a global project, but its context is carefully described.
	ii	Clear and relevant description (where appropriate) of the status and needs of the target groups for the intervention	Yes	The status and needs of the target groups are described, including those elements that could not be addressed by the project.
Question 4.		Are key stakeholders and their contributions clearly identified?	100%	
	i	Identification of implementing agency(ies), development partners, primary duty bearers, secondary duty bearers, and rights holders	Yes	This is a joint project, with many national partners. They are well described.
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF	Yes	The roles are well described.
SECTION C:		**EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)**	100%	Comments on Rating (include explanations for any criterion not rated)
Question 5.		Is the purpose of the evaluation clearly described?	100%	
	i	Specific identification of how the evaluation is intended to be used and to what this use is expected to achieve	Yes	The evaluation is to be used in the final year and in successive plans.
	ii	Identification of appropriate primary intended users of the evaluation	Yes	The intended users are identified.
Question 6.		Are the objectives and scope of the evaluation clear and realistic?	100%	
	i	Clear and complete description of what the evaluation seeks to achieve by the end of the process with reference to any changes made to the objectives included in the ToR (if applicable)	Yes	The evaluators are clear about the intended results of the evaluation.
	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 is defined, including why and how only three of the 12 countries were included in field visits based on time and resources issues.
Question 7.		Is the results chain or logic well articulated?	100%	
	i	Clear description of the intervention's intended results, or of the parts of the results chain that are applicable to the evaluation	Yes	The results chain is carefully described, including those elements that could not be measured.
	ii	The causal relationship between outputs and outcomes is presented in narrative and/or graphic form (theory of change, logic model, results chain, evaluation matrix, etc.)	Yes	When there was a clear causal connection to be measured, this was indicated. Some elements could not be causally measured.
SECTION D:		**EVALUATION METHODOLOGY (weight 20%)**	93%	Comments on Rating (include explanations for any criterion not rated)
Question 8.		Does the evaluation provide a relevant list of evaluation criteria that are explicitly justified as appropriate for the purpose of the evaluation? UNICEF evaluation standards refer to the OECD/DAC criteria. Not all OECD/DAC criteria are relevant to all evaluation objectives and scopes. Standard OECD DAC Criteria include: Relevance; Effectiveness; Efficiency; Sustainability; Impact. Evaluations should also consider equity, gender and human rights (these can be mainstreamed into other criteria). Humanitarian evaluations should also consider Coverage; Connectedness; Coordination; Protection; Security.	100%	
	i	Clear and relevant presentation of the evaluation framework including clear evaluation questions used to guide the evaluation	Yes	The eight evaluation questions (in four categories) were well described in Annex A and were used to guide the evaluation.
Question 9.		Does the report specify methods for data collection, analysis, and sampling?	88%	
	i	Clear and complete description of a relevant design and set of methods that are suitable for the evaluation's purpose, objectives and scope	Yes	The design and methods are appropriate for the project and its context.

ii	Clear and complete description of the data sources, rationale for their selection and sampling strategy. This should include a description of how diverse perspectives are captured (or if not, provide reasons for this), how accuracy is ensured, and the extent to which data limitations are mitigated.	Partially		For documents and in-depth interviews, the choices were clearly defined. The selection of countries was also defined by criteria. Within the countries, however, why particular villages were chosen for the focus groups was not clearly indicated.
iii	Clear and complete description of the methods of analysis, including triangulation of multiple lines and levels of evidence (if relevant)	Yes		The methods of analysis are clearly described.
iv	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		The limitations were set out clearly by the evaluators as well as steps taken to mitigate.
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:	100%		
i	Explicit reference to the obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability)	Yes		UNEG guidelines were cited and referenced.
ii	Description of ethical safeguards for participants appropriate for the issues described (respect for dignity and diversity, right to self-determination, fair representation, compliance with codes for vulnerable groups, confidentiality, and avoidance of harm). For those cases where the evaluation involves interviewing children, explicit reference is made to the UNICEF procedures for Ethical Research Involving Children	Yes		A text box was used to highlight the teams adherence to ethical standards. UNEG guidelines were referenced, and actions taken within this evaluation were well described. This is notable example of good practice.
SECTION E:	**EVALUATION FINDINGS (weight 25%)**	93%		Comments on Rating (include explanations for any criterion not rated)
Question 11.	Do the findings clearly address all evaluation objectives and scope?	100%		
i	Findings marshal sufficient levels of evidence to systematically address all of the evaluation's questions and criteria	Yes		The mobilization of evidence is impressive given the evaluation questions. There was solid use of document analysis to demonstrate global patterns, but also field visits that allowed detailed presentation of results.
ii	Reference to the intervention's results framework in the formulation of the findings	Yes		The findings are organized by outcomes and outputs.
Question 12.	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.	83%		
i	The evaluation clearly presents multiple lines (including multiple time series) and levels (output, outcome, and appropriate disaggregation) of credible evidence.	Yes		Depending on the issue, there are multiple lines of evidence.
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.	Yes		The findings are supported by the evidence presented. In particular the evaluators use the carefully-defined theory of change as a basis for assessing results, both positive and negative. For example, the finding on whether "GPECM reaches the most vulnerable" states that "While the programme has made strategic choices in targeting areas with high child marriage prevalence, interventions supported by the GPECM have yet to reach some of the most vulnerable including those in less accessible and remote areas who are among the furthest behind."
iii	The causal factors (contextual, organizational, managerial, etc.) leading to achievement or non-achievement of results are clearly identified. For theory-based evaluations, findings analyse the logical chain (progression -or not- from implementation to results).	Partially		In many cases, there is a clear identification of causal factors. In a few, where the causal connection between the results and the project is not clear, the evidence was less compelling. For example, in Finding 10 relating to Outcome 5 that "Government(s) and partners within and across countries support and promote the generation and use of robust data and evidence to inform programme design, track progress and document lessons," the evaluation finds that "PECM investments in research and data have contributed to building a stronger evidence base on child marriage. Tracking has not offered an indication on the quality and applicability of data generated; some research has been more strategic than other research. Few studies have offered any indication on the impact of interventions." The causal connection between the research and the design is not clear.
Question 13.	Does the evaluation assess and use the intervention's Results Based Management elements?	100%		

	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)	Yes	Considerable effort was made to examine the adequacy of the monitoring system.
	ii	Assessment of the use of monitoring data in decision making	Yes	There is a specific section of the findings addressing this.
SECTION F:		**EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 10%)**	88%	Comments on Rating (include explanations for any criterion not rated)
Question 14.		Do the conclusions present an objective overall assessment of the intervention?	100%	
	i	Clear and complete description of the strengths and weaknesses of the intervention that adds insight and analysis beyond the findings	Yes	Both the strengths and weaknesses are referenced in the conclusions, but especially the strengths.
	ii	Description of the foreseeable implications of the findings for the future of the intervention (if formative evaluation or if the implementation is expected to continue or have additional phase)	Yes	Suggestions about the future are a major part of the conclusions.
	iii	The conclusions are derived appropriately from findings	Yes	They are all derived from the very detailed findings.
Question 15.		Are logical and informative lessons learned identified?	50%	
	i	Identified lessons that stem logically from the findings, presents an analysis of how they can be applied to different contexts and/or different sectors, and takes into account evidential limitations such as generalizing from single point observations.	Partially	There is no specific section on lessons learned, but many of the conclusions have a more general application.
SECTION G:		**RECOMMENDATIONS (weight 15%)**	100%	Comments on Rating (include explanations for any criterion not rated)
Question 16.		Are recommendations well grounded in the evaluation?	100%	
	i	Recommendations are logically derived from the findings and/or conclusions	Yes	The findings are clearly the basis for the recommendations.
	ii	Recommendations are useful to primary intended users and uses (relevant to the intervention and provide realistic description of how they can be made operational in the context of the evaluation)	Yes	The recommendations are for global, regional and country offices, depending on the focus.
	iii	Clear description of the process for developing recommendations, including a relevant explanation if the level of participation of stakeholders at this stage is not in proportion with the level of participation in the intervention and/or in the conduct of the evaluation	Yes	The process of development of recommendations with participation was specified in the ToR and implemented as required, especially via discussions with the Evaluation Management Group.
Question 17.		Are recommendations clearly presented?	100%	
	i	Clear identification of target group for action for each recommendation (or clearly clustered group of recommendations)	Yes	The target group in each case was specified.
	ii	Clear prioritization and/or classification of recommendations to support use.	Yes	The recommendations were organized by subject matter.
SECTION H:		**EVALUATION STRUCTURE/PRESENTATION (weight 5%)**	100%	Comments on Rating (include explanations for any criterion not rated)
Question 18.		Does the evaluation report include all relevant information?	100%	
	i	Opening pages include: Name of evaluated object, timeframe of the evaluation, date of report, location of evaluated object, names and/or organization(s) of the evaluator(s), name of organization commissioning the evaluation, table of contents -including, as relevant, tables, graphs, figures, annexes-; list of acronyms/abbreviations, page numbers	Yes	The pages are as desired.
	ii	Annexes should include, when not present in the body of the report: Terms of Reference, Evaluation matrix, list of interviewees, 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, copy of the results chain, information about the evaluator(s)	Yes	The annexes are complete.
Question 19.		Is the report logically structured?	100%	
	i	The structure is easy to identify and navigate (for instance, with numbered sections, clear titles and sub-titles)	Yes	The structure is easy to navigate.
	ii	Context, purpose and methodology would normally precede findings, which would normally be followed by conclusions, lessons learned and recommendations	Yes	The order is as desired.
Question 20.		Is the report well presented?	100%	
	i	Report is easy to understand (written in accessible way for intended audience) and generally free from grammar, spelling and punctuation errors	Yes	The report is well-drafted.
SECTION I:		**EVALUATION PRINCIPLES (weight 10%)**	96%	Comments on Rating (include explanations for any criterion not rated)
Question 21.		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?	100%	
	i	Reference and use of rights-based framework, and/or CRC, and/or CCC, and/or CEDAW and/or other rights related benchmarks in the design of the evaluation	Yes	There is clear reference to the rights-based framework (which includes a large number of benchmarks).

	ii	Clear description of the level of participation of key stakeholders in the conduct of the evaluation, (for example, a reference group is established, stakeholders are involved as informants or in data gathering)	Yes	The role of the Evaluation Management Group is clear and involved the main stakeholders. Also, in country visits, stakeholders were informants providing data.
	iii	Stylistic evidence of the inclusion of these considerations which can include: using human-rights language; gender-sensitive and child-sensitive writing; disaggregating data by gender, age and disability groups; disaggregating data by socially excluded groups.	Yes	The considerations were included in the evaluation at all points.
Question 22.		Does the evaluation assess the extent to which the implementation of the intervention addressed gender, equity & child rights?	100%	
	i	Identification and assessment of the presence or absence of equity considerations in the design and implementation of the intervention (such as the involvement in the intervention of right holders, duty bearers, and socially marginalized groups, and the differential benefits received by different groups of children)	Yes	Equity considerations were built into the evaluation and reflected in findings and conclusions.
	ii	Clear proportionality between the level of participation in the intervention and in the evaluation, or clear explanation of deviation from this principle (this may be related to specifications of the ToRs, inaccessibility of stakeholders at the time of the evaluation, budgetary constraints, etc.)	Yes	The principles were followed in that those who were part of the intervention were providers of information about results.
Question 23.		Does the evaluation meet UN SWAP evaluation performance indicators? (Note: this question will be rated according to UN SWAP standards)	89%	
	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.	Fully integrated	The project is a GEEW-focused project, and this was clearly built into the evaluation. Gender issues are integrated into the evaluation questions and indicators. One issue of importance was the need to look at boys and men as well as women in examining gender issues.
	ii	A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Satisfactorily integrated	The methodology did not explicitly address how evaluation design and implementation was gender responsive; evaluation participants were not gender-disaggregated in this section although they were in the annex. Otherwise, the methodology appeared to be gender responsive throughout with a good diversity of stakeholders consulted, ethical standards upheld, and gender-disaggregated data presented throughout the findings.
	iii	The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Fully integrated	The findings include a specific section on gender and inclusivity. In addition to ensuring that the project is more effective (in part by better country-level planning and generally in terms of improved monitoring and evaluation, as well as SDG 5.3), the need to include boys and men as well as girls and women in the project is emphasized. These issues are reflected in the background section and carried through into the recommendations.

	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?
	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?