GEROS Evaluation Quality Assurance Tool

Version: September 2021

UNFPA-UNICEF Global Programme to End Child Marriage, 2020-2023

REPORT RATING SUMMARY		
Overall Rating	92%	Highly Satisfactory
● ● ● ● ● Exceptional (96% - 100%)	5	
● ● ● ● Highly Satisfactory (87.5% - 95.99%)	**4**	**Exceeds UNICEF/UNEG standards for evaluation reports.** **Decision makers may use the evaluation with a high degree of confidence**
● ● ● ▬ Satisfactory (62.5% - 87.49%)	3	
● ● ▬ ▬ Fair (35% - 62.49%)	2	
● ▬ ▬ ▬ Unsatisfactory (0% - 34.99%)	1	

REPORT DETAILS	
Title of the evaluation report	UNFPA-UNICEF Global Programme to End Child Marriage, 2020-2023
Report sequence number	Evaluation Office/65/2023/19915
Region	HQ
Year of report	2023
Office	Evaluation Office
Coverage (countries)	Global
ToRs present	Yes
Date of review (dd/mmm/yyyy)	March 4, 2024
Name of review firm	IOD PARC

CLASSIFICATION OF EVALUATION REPORT	
Management of evaluation *(Managerial control and oversight of evaluation*	Jointly managed with one or more UN agencies
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	No
Every child has an equitable chance in life	Yes
Gender equality (cross-cutting)	Yes
Humanitarian action (cross-cutting)	No
Evaluation object	Joint Programme
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 *(number)*	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.*

The Global Programme to End Child Marriage is a complex joint agency long term programme between UNICEF and UNFPA which started in 2016 and concludes in 2030, and which operates in 12 countries.   This evaluation of Phase 2 builds on the Phase 1 evaluation, as well as clearly situating itself in wider research and learning to support the programme as it moves into Phase 3. The evaluation is of very high quality, and meets almost all UNICEF criteria for an excellent evaluation which can be used as an example of best practice. Key areas of excellence include:

• The attention to evaluating gender transformative approaches, and to gender intersectionality as an approach and a method of analysis.
• The methodology is very strong, with multiple methods and tools deployed including desk review, country visits, deep dives, KIIs with stakeholders and crucially FGDs with parents and other adults, as well as youth consultations with children and young people. The sampling strategy was designed to operate at global and local levels to ensure diversity of experiences and voices were present in the findings, and innovative methods of analysis applied.   Data collection tools are provided as an Annex and are both detailed and robust. Overall the design and methods are highly relevant, robust and clearly described, taking different contexts of this global evaluation into account
• The findings section is systematic and clear, using multiple sources of data for triangulation purposes, as well as different analytical lens, for example using a realist lens to address questions relating to effectiveness. The Programme Theory of Change is a guiding document and is used, and assessed throughout the findings section.
• Conclusions are clear and add additional relevance, with good traceability to the evidence and findings to which they relate. Lessons are succinct and relevant, with much wider applicability, particularly for programmes which aim to take a gendered transformative approach.

While this evaluation is of a very high standard there are several (mostly) minor areas to consider for improvement, these are:
• Ideally the executive summary should include methodology, conclusions, and lessons learned. The findings and recommendations could be shorted and the executive summary restructured to include all relevant sections.
•Describing how the programme works in practice, in different country contexts, and with different rightsholders would aid in understanding how the programme operates on the ground, as well as what is being evaluated. Appreciating that much of this is present in the findings section, it would be helpful to include a short summary in the context, with links to more detail in an annex.
• Minor editing issues to be reviewed, such as the ToC (Annex 2) being blurry and page numbers to start at the introduction, and appear on every page.
• Acknowledging that the report reads well, is written clearly and logically, and represents findings from global to local level (in 12 countries), the ToR stated that the final report should be 60 pages (excluding annexes) this final report is 113 pages (although it is noted the first 15 pages could be excluded).
• The ERG membership and their role could be more clearly outlined/ included as an Annex.
• The process for developing the recommendations should be clearly described. *(to note, if this were included the evaluation would be rated as exceptional, as this is the only question which was not addressed at all/was rated as "no")*

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.*

This evaluation is of a very high standard and meets almost all of UNICEFs criteria for a good evaluation. Areas for improvement include:

• Ideally the executive summary would include a short outline of the methodology used, as well as conclusions. To include this withing the 5 page limit, the findings could be shortened or merged with conclusions to present a more well rounded summary that reflects all key sections of the evaluation report. Recommendations could also be presented more succinctly.
• To aid in understanding, consider explaining how the programme works in practice, at country level, and the number of beneficiaries/ rightsholders reached as part of the context section, not only in findings.
• At just under 100 pages in length (for the actual report, excluding the executive summary, annexes are separate) the report is long, however, it does evaluate a global programme that is implemented in 12 countries, is long term (15 years in length) and a joint programme. It is also a very well structured and written report. However, the ToR noted that the final evaluation report should be 60 pages long. It may be sensible to consider if that page limit was appropriate for a programme evaluation of this nature, or if there were details of the report (elements of the findings section for example) which could have been shortened.
• It would be useful to include detail about, and membership of the Evaluation Reference Group in the conduct of this evaluation, possibly as an Annex.
• The process for developing the recommendations should be described.

SECTION RATINGS			
SECTION A:	**EXECUTIVE SUMMARY (weight 5%)**	67%	Comments on Rating
Question 1.	Can the executive summary inform decision-making?		
i	Is clearly presented, serves as a standalone document useful for informing decision making, and is of relevant conciseness and depth for key users (Maximum of 5 pages unless otherwise specified in ToR).	Partially	The executive summary is clearly presented and serves as a standalone document. It is just over 5 pages long, although no page limit was given in the ToRs. The findings section, could be shortened and the overall executive summary slightly restructured to stay within the 5 page limit.
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 executive summary includes a good context section, but the findings are overlong, and there are no outline of the methodology, conclusions or lessons learned. It would be suggested that the findings section is shortened, or a merged section of conclusions and findings drafted, along with a short list of the key lessons. Recommendations could be shorted (they currently span 3 pages), and a methodology paragraph included.
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	All significant information is included, even with the lack of a conclusions section the findings section does a good job of explaining the findings, and may be better referred as findings and conclusions. No new information is presented.
SECTION B:	**BACKGROUND (weight 5%)**	93%	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.	Yes	The object of this evaluation is The Global Programme to End Child Marriage (GPECM) - a joint initiative between UNICEF and UNFPA "that turns United Nations commitments into tangible actions for children, supporting Governments and civil society partners to accelerate actions to end child marriage." (pp 20). The intervention is described well in relation to: • Location - 12 countries in five regions • Timeline - the intervention started in 2016 and concludes in 2030, with three phases of implementation. • Implementation status - the intervention is at the end of phase 2. Table 1 provides a very useful and succinct summary of each phase • Cost/ Budget - The proposed and agreed costs for Phase 2 are outlined, with details of the impact COVID 19 had on programme spend. Overall the funding and governance structures are clear.
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/cites, sub-regions…) and in terms of numbers reached, with disaggregation by gender, age, disability . . . (as appropriate to the purpose of the evaluation).	Partially	There is a clear and relevant description of intended rights holders, with girls aged 10-19 years identified as the primary rights holders, and secondary rights holders are women, men's, boys and girls, and wider families and communities who benefit from the programme, or from its wider reach. The rightsholders are located in the following regions and countries: • Eastern and Southern Africa: Ethiopia, Mozambique, Uganda and Zambia • Middle East and North Africa: Yemen • West and Central Africa: Burkina Faso, Ghana, Niger and Sierra Leone • South Asia: Bangladesh, India and Nepal Duty bearers are described as government who have made a commitment to end child marriage, i.e. the governments of the 12 programme countries. Other duty bearers include international, national and local organisations. Greater clarity could be provided on the specific roles of different types of duty bearer. This section could have been improved if there was detail in the context section about how the programme operates in practice, and numbers of people reached by the programme in each country to date, disaggregated by age, gender or other characteristics. Appreciating that the findings section does contain some of this information it may have been useful to summarise this in the context/ introduction sections, and perhaps include an annex with more detail to aid in understanding how the programme operates in practice, in different contexts, with different people (girls, boys, women and men).
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 clear and relevant description of the context of the intervention, including a definition of child marriage, and the negative impact of this, especially for girls. Several infographics (Figure 1 and 2) usefully illustrate the scale of the issue, and change over time (in different regions). The cultural, power and socioeconomic contextual factors, are well explained. There is also good attention given to the global factors and history of global action against child marriage, following a historical thread of declarations and treaties. Accompanying this is a clear outline of the institutional factors, policies and practice, within UNICEF and UNFPA, with direct relevance to the intervention.
	ii	Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Yes	There is direct reference, and strong linkages made to SDG 5, and specifically target 5.3.
	iii	Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Yes	There is a clear and relevant description at the global level of the status and needs of those subject to child marriage, and who are the rightsholders/ beneficiaries of this intervention.
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).	Yes	A steering committee oversees implementation of the programme and is comprised of a representative from UNICEF, UNFPA, donors and a government representative from each programme country. The Steering committee is support by a partners advisory group, which provides technical advice and serves as a knowledge sharing platform. This is comprised of United Nations agencies, donors and civil society organizations (CSOs). There is also a global programme support unit, and technical advisory groups. While high level, the role of stakeholders is clearly described in relation to the programme, and to each other.
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Yes	Specific management/ governance contributions of UNICEF and UNFPA are well defined and described. Donors are listed, and financial contributions (planned and actual) are noted.
SECTION C:		**EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)**	100%	Comments on Rating
Question 5.		Is the purpose of the evaluation clearly described?		
	i	Purpose of evaluation is clearly defined, including why it was needed at that point in time, its intended use, and key intended users.	Yes	The purpose of this evaluation is clearly defined as being to independently assess the achievement of outputs and outcomes in phase II, provide information that supports evidence-based decision-making, and inform the work of the GPECM beyond 2023. It is clear why it is needed at this point at the end of Phase 2, to inform the implementation of Phase 3.
Question 6.		Are the objectives and scope of the evaluation clear and realistic?		
	i	Clear and complete description of what the evaluation seeks to achieve by the end of the process with reference to any changes made to the objectives included in the ToR (if applicable).	Yes	There is a clear and complete description of what the evaluation seeks to achieve, and it is noted that there are no changes to the objectives outlined in the ToR.
	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	There is a clear and relevant description of the scope of the evaluation in relation to: • Thematic/ Analytical Scope - the evaluation assessed the programme outcomes, thematic areas outlined in the evaluation matrix, and specific OECD DAC criteria. it also tested the assumptions of the ToC. • Chronologically - the evaluation covered the second phase of the GPECM from January 2020 until the completion of data collection during the second quarter of 2023 • Geographically - The evaluation covered the results of the programme and assessed its contribution to ending child marriage at the global, regional and country levels. Of the 12 programme countries, four countries; Ethiopia, Ghana, India and Sierra Leone were selected for in-country deep dive case studies. One area which was suggested for investigation in the ToR, relating to Value for money, was not included as the evaluation determined this was out side the scope of this evaluation. This is clearly explained.
Question 7.		Is the theory of change, results chain or logic well articulated?		
	i	Clear description of the intervention's intended results, or of the parts of the results chain that are applicable to, or are being tested by, the evaluation.	Yes	The report is clear that phase 2 of the programme is being evaluated, but provided further detail on specific focus area of this evaluation, including • Gender-transformative approaches (GTA) and how well the programme understands the GTA and its implementation • Leave No-one behind, and a focus on reaching how well the intervention reached the marginalized girls • Other areas including catalytic role of the programme, added value of joint programming, and adaptation to crisis. These are clear, and well described, adapting to changing contexts within the overarching programme, and the world (e.g. Covid 19). However, as noted earlier a description of the programme at the different levels it operate would be useful to help understand these focus area from a practical and country perspective in particular.

	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	Causal relationships between outputs and outcomes are presented in narrative form throughout the report, and graphically represented in the ToC and evaluation matrix (Annexes 2 and 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	The global Theory of Change for the programme was assessed as part of the Phase 1 evaluation, and was revised for phase II. This evaluation report assesses and tests the Phase 2 ToC (provided at Annex 2). A minor point to note is that the ToC in Annex 2 is blurred and difficult to read, presumably because it is a copied image, as such it would be recommended to include a clearer version of this.
SECTION D:		**EVALUATION DESIGN AND METHODOLOGY (weight 20%)**	95%	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? *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	Evaluation questions and sub-questions are appropriate for meeting the objectives and purpose of the evaluation. The relevant criteria are specified and are aligned with the questions.	Yes	Evaluation questions and sub questions are appropriate for meeting the objectives and purpose of the evaluation. Relevant criteria are spelled out, not only in the evaluation matrix, but in the main body of the report, along with different methodologies which would apply best for different criteria. This is excellent practice showing a nuanced understanding of OECD DAC criteria, and their application and utility in meeting the objectives of this evaluation.
	ii	In addition to the questions and sub-questions, the evaluation matrix includes indicators, benchmarks, assumptions and/or other processes from which the analysis can be based and conclusions drawn.	Yes	The evaluation matrix is excellent, and for each sub question provides a list of the assumptions to be assessed, the Indicators/markers, the sources of information and the methods for data collection.
Question 9.		Does the report specify adequate methods for data collection, analysis, and sampling?		
	i	Evaluation design and set of methods are relevant and adequately robust for the evaluation's purpose, objectives and scope; and are fully and clearly described.	Yes	The evaluation used a theory-based, mixed methods and a multiple case-study design. However it is well noted that to test certain elements of the ToC contribution analysis and realist approaches have been applied, with these methods particularly appropriate for assessing effectiveness in different contexts/ countries. The methodology provides a good example of how different approaches can be deployed to address different questions or criteria within an evaluation. Human right based approaches and a gender lens were also applied across the evaluation as whole, and clearly explained. Figure 4 (pp 30) usefully provides an overview of methods and tools, which are also well described narratively. Further detail on design and methods is provided in Annex 4, while Annex 5 contains robust data collection tools. Overall the design and methods are highly relevant, robust and clearly described, taking different contexts of this global evaluation into account.
	ii	Data sources are appropriate - these would normally include qualitative and quantitative sources (unless otherwise specified in the ToR) - and are all clearly described.	Yes	Data sources are appropriate, and include qualitative and quantitative sources. All are clearly described.
	iii	Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or, if not, provide reasons for this).	Yes	Sampling strategies are provided, and include: •Geographical sampling for selection of case study countries including criteria such as countries/ regions, size and complexity of programme, budget and expenditure, humanitarian context, the implementation of GTA activities, number of past studies and evaluations, and political constraints. •Deep dive sampling where convenience sampling was used for the selection of locations to visit within deep dive countries. Criteria included both UNICEF and UNFPA programming, across a variety of interventions at least two geographic locations (where possible) providing diversity in terms of culture, religion, language etc (pp 34). Security concerns were also considered. •Selection of respondents (KIIs, FGDs, survey) - "Sampling was purposeful given the time and aims of the evaluation. Nonetheless, the sampling strategy focused on ensuring diversity of stakeholder groups with a view to capturing multiple viewpoints and experiences." (pp34). Overall it is clear that the sampling strategy was designed to operate from the global to the local levels, and to ensure diversity of experiences and voices was included wherever possible.
	iv	Clear and complete description of the methods of analysis.	Yes	The methods of analysis are clear well described, including the process of coding, use of analytical frameworks, and of extracting, analysing, and triangulating data, applying a realist and contribution analysis lens. Inductive and deductive methos were used to ensure all intended (or unintended) outcomes were captured and analysed.

	v	Methodology allows for drawing causal connections between outputs and expected outcomes.	Yes	The methodology is excellent, and allows for the drawing of causal connections between outputs and outcomes in different contexts.
	vi	Clear and complete description of limitations and constraints faced by the evaluation, including gaps in the evidence that was generated and mitigation of bias, and how these were addressed by the evaluators (as feasible).	Partially	There is a clear section on limitations (pp 35). While this considers the main issues (access to people and data), mitigations were only provided for access to data, it is assumed that if the most marginalised were the hardest to reach then there was no mitigation against this. This section could be improved if there was confirmation that there was no mitigation against lack of access to marginalised people, and if bias were present, and mitigation of it were discusses here.
Question 10.		Are ethical issues and considerations described? The evaluation should be guided by the UNEG ethical standards for evaluation. As such, the evaluation report should include:		
	i	Explicit and contextualized reference to the UNEG obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability) and/or the principles in the 2020 revised UNEG Ethical Guidelines (integrity, accountability, respect, beneficence).	Yes	There is explicit and contextualised reference to the UNEG obligations and the robust adherence to the principles in the 2020 revised UNEG ethical guidance.
	ii	Description of ethical safeguards for participants appropriate for the issues relevant to methodology and how they are applied (respect for dignity and diversity, right to self-determination, fair representation, compliance with codes for vulnerable groups, confidentiality, and avoidance of harm). For those cases where the evaluation involves interviewing children, explicit reference is made to the UNICEF procedures for Ethical Research Involving Children.	Yes	Ethical safeguards are well described, and appropriate, including respect and consideration for dignity and diversity. There is also explicit reference to global guidelines in working with youth and children including that the: "evaluation team adhered to global guidelines, including the UNICEF Guidance Note on Adolescent Participation in UNICEF Monitoring and Evaluation, the Minimum Standards for Child Protection in Humanitarian Action, which includes ethical guidelines on research and interaction with children and adolescents, and child safeguarding protocols established by UNICEF." (pp35). Additional consideration is given to ensuring voices of marginalised people, and those of children were heard in the most appropriate, safe and ethical way, with a protocol developed for each country including a referral process for any issues of concern raised during interviews, consultations of focus groups. Annex 7 contains further detail on ethical safeguards including consent forms and referral pathways.
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 evident in several ways such as the design of the methodology (i.e. use of technology for data gathering, extensive participatory processes, systematic analysis processes such as collaborative outcomes reporting and incorporation of big data, specific strategies to address complexity such as outcome harvesting, strong child rights focus), or ways of sharing of evaluation results.	Yes	The application of a realist approach in the design and analysis of this evaluation is considered innovative, and careful thought has been given as to how to do this appropriately. As such the evaluation team selected only questions related to effectiveness, as a realist lens could only be applied in addressing those questions. This shows a nuanced and deep understanding of realist methodologies, and how their application can be used to best effect. The evaluation methodology also manages complexity very well.
SECTION E:		EVALUATION FINDINGS (weight 25%)	100%	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.	Yes	Findings marshal significant levels of evidence, to systematically address the evaluations questions and sub questions, taking account of the criteria to which they relate. The findings section is structured clearly around evaluation questions, noting which criteria they relate to, and identifying findings within that. Evidence is provided from a sample of all countries the programme operates in, and there is good use of contextualised evidence and analysis for different country contexts.
	ii	Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Yes	There is explicit use of the programme ToC in the formation of the findings, both where this is validated by the findings, or where evidence suggests there is still further to go, for example Finding 14 (pp 77), notes the key links in the ToC between socioeconomic empowerment, and social protection and higher incidences of child marriage. However, there is still a lack of country programmes, and government support (in some countries) that directly look to address child marriage by improving girls (and families) socio economic outlook, indicating further work is needed in this area. Another clear example of linking assessment of the intervention activities to the ToC outcomes is provided in addressing Finding 16, Table 10 (pp87), which clearly outlines activities undertaken by the programme, their relative success, and which outcome of the ToC they directly relate to.
Question 13.		Are evaluation findings derived from the conscientious, explicit and judicious use of the best available, objective, reliable and valid data and by accurate quantitative and qualitative analysis of evidence.		
	i	Evaluation uses credible forms of qualitative and quantitative data. It presents both output and outcome-level data as relevant to the evaluation framework. Triangulation is evident through the use of multiple data sources.	Yes	The evaluation uses credible forms of qualitative and quantitative data, and presets both output and outcome level data (where available). Triangulation is evident and there is strong evidence of the use of multiple and varied data sources.

	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.	Yes	Findings are clearly supported by, and responds to, the evidence presented, both positive and negative. The evaluation also uses multiple ways of analysing and comparing results/ performance, or example Finding 13 on gender transformative approaches outlines the indicators or benchmarks that were used by different countries, accepting that all countries are at a different stage in the journey towards gender equality. Acknowledging this the programme designed a gender transformative strategy, with priorities, aims and objectives, and indicators of success to be monitored. The report uses this as one means of assessing progress towards gender transformative approaches, along with other data from KIIs, desk review etc, to analyse findings from multiple perspectives, including local, national and international benchmarks (where feasible and appropriate).
	iii	Causal factors (contextual, organizational, managerial, etc.) leading to achievement or non-achievement of results are clearly identified. For theory-based evaluations, findings also analyse the logical chain (progression -or not- from implementation to results).	Yes	Causal factors are clearly identified, including different country contexts, as well as management of the programme, and role of partners/ CSOs. The Theory of Change is references throughout with regard to testing assumptions for the Phase 2 ToC, as well as identifying where implementation had progressed to outputs, outcomes, and results.
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	The evaluation assesses the adequacy of the programmes monitoring system overall, and in detail in Finding 19 (pp 90). It concludes that output level data is being captured well, however, outcome level monitoring and results relating to this could be improved. Recommendations from the phase 1 evaluation are referenced, however it is noted that this did not result in significant improvements to the programme monitoring, excepting in the Ethiopian context. Tools used in some countries are discussed and their limitations discussed. The need for appropriate gender responsive monitoring and tools, as well as training to support this is well noted.
SECTION F:		EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 10%)	100%	Comments on Rating
Question 15.		Do the conclusions clearly present an objective overall assessment of the intervention?		
	i	Conclusions are clearly formulated and reflect the purpose and objectives of the evaluation. They are sufficiently forward looking (if a formative evaluation or if the implementation is expected to continue or have additional phase).	Yes	Conclusions are clearly formulated and reflect the purpose and objectives of the evaluation. They are situated within the context of this long term programme, but also sufficiently forward looking to inform the next phase of its implementation.
	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	Conclusions are derived appropriately from findings, and usefully under the heading for each conclusion is a note of the findings which this conclusions relates to, which is good practice for traceability of findings and conclusions. The conclusions add insight and analysis beyond the findings, presenting a contextualised and realist assessment of the strengths and limitations of the programme as it moves into its third and final phase.
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.	Yes	Identified lessons stem logically from the findings, with the first finding focusing on gender transformative approaches, outlined in the methodology, and analysed in the findings and conclusions. The other 4 lessons also stem logically from the findings relating to role of national laws and policies, joint programming, targeting of marginalised/ at risk girls in programming and convergent programming to support intersectional approaches. All lessons have wider applicability and relevance beyond the object of the evaluation
	ii	Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Yes	Lessons are clearly and concisely presented and have sufficient detail to be useful for the intended audience as well as any programmes which aim to take a gender transformative approach, and engage meaningfully with issues of intersectionality and gender in general, as well as child marriage explicitly.
SECTION G:		RECOMMENDATIONS (weight 15%)	75%	Comments on Rating
Question 17.		Are recommendations well grounded in the evaluation?		
	i	Recommendations align with the evaluation purpose, are clearly formulated and logically derived from the findings and/or conclusions.	Yes	Recommendations align with the evaluation purpose, are clearly formulated and logically derive from the conclusions, with a note of which conclusions relates to which recommendation. As the conclusions are directly linked to the individual findings there is excellent ability to trace the evidence in this evaluation.
	ii	Recommendations are useful and actionable for primary intended users and uses (relevant to the intervention); guidance is given for implementation, as appropriate.	Yes	Recommendations are useful and actionable for primary intended users, and guidance is given for implementation.
	iii	Process for developing the recommendations is described, and includes the Involvement of duty-bearers, as well as rights holders when feasible (or explanation given for why they were not involved).	No	The process for developing the recommendations is not described.
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	There is clear identification of duty bearers for implementing each recommendations, with all recommendations applying to both UNICEF and UNFPA, to be actioned by the 12 country offices of each organisation where the programme is being implemented. HQ and regional offices in both organisations are to support their implementation. Recommendations are clearly prioritised, with close attention to the fact the recommendations are to inform the next phase of the programme. A clear prioritization and timeline is provided for each recommendation.
SECTION H:		**REPORT STRUCTURE AND PRESENTATION (weight 5%)**	92%	Comments on Rating
Question 19.		Does the evaluation report include all relevant information?		
	i	Opening pages include: Name of evaluated object, timeframe of the evaluation, date of report, location of evaluated object, names and/or organization(s) of the evaluator(s), name of organization commissioning the evaluation, table of contents -including, as relevant, tables, graphs, figures, annexes; list of acronyms/abbreviations, page numbers.	Yes	The opening pages contain all relevant and expected information for a UNICEF evaluation. A minor point to note that page numbers should start on the first page of the report introduction and not include the contents and executive summary sections. Also page number sonly appear on every second page.
	ii	Annexes include: terms of reference, evaluation matrix, list of interviewees, results chain/ToC/logical framework (unless included in report body), list of site visits, data collection instruments (such as survey or interview questionnaires), list of documentary evidence. Other appropriate annexes could include: additional details on methodology, information about the evaluator(s), etc.	Yes	The annexes are thorough and add to the depth and clarity of the overall report, They include ToRs, evaluation matrix, field visits conducted, stakeholders consulted, data collection tools, details on methodology and ethics.
Question 20.		Is the report logically structured?		
	i	Structure is easy to identify and navigate (for instance, with numbered sections, clear titles and sub-titles, well formatted).	Yes	The structure is excellent, the report flows logically from one section to the next, with good use of headings and sub headings, with clearly named and numbered sections. It is very well formatted.
	ii	Structure accords to UNICEF guidelines for evaluation reports: context, purpose and methodology would normally precede findings, which would normally be followed by conclusions, lessons learned and recommendations.	Yes	The structure accords closely to UNICEF guidelines for evaluation reports.
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.	Partially	The report is 113 pages long, but approximately 15 of those should not be included as they include contents page and executive summary. At almost 100 pages for the main report this is still overlong, particularly when the ToR noted that the expectation was that the report should be 60 pages in length. However, it should be noted that the report reads well, contains excellent and relevant information and does a good job of addressing findings across 12 countries, as well addressing the global and joint programme context.
	ii	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 exceptionally well written, clear and accessible for intended audience, with a strong narrative running though the report linking each section logically and systematically. It is free from grammar, spelling and punctuation errors.
	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	There is frequent use of visual aids - which is essential and helps to explain data and information about this complex programme. All visual aids are clearly presented, labelled and referenced in the text.
SECTION I:		**EVALUATION PRINCIPLES (weight 10%)**	93%	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	The evaluation firmly takes a human rights based approach, with the child marriage clearly framed as against basic human rights. One key aspect of the programme being evaluated is its ability to influence national and international policies, including recognition of the human rights of children not to be subject to forced marriage. This HRBA approach was integrated from the inception of the programme, and formalised in the Theory of change, subsequent programming, and evaluation of it. Other rights based frameworks are referenced, including CEDAW.
	ii	Clear description of the level of participation of key rights holders and duty bearers in the conduct of the evaluation (for example, a reference group is established, stakeholders are involved as informants or in data gathering).	Partially	The evaluation reference groups is mentioned in several sections of the report, noting where they were consulted in determining the scope of the evaluation development of the evaluation matrix and their role in quality assurance. However, the ERG and their role is not clearly defined. Ideally an annex should be included noting membership of the ERG, and their role.
	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 both empowering and inclusive and aims to provide cultural context as well as providing data and evidence across, and between marginalised groups. Given the intersectional nature of gender inequality, consideration was given to ensure a balance of respondents in terms of gender, disability, socioeconomic status, ethnicity, religious affiliation, and other factors that create marginalization, disenfranchisement, or vulnerability.

	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	The report notes that the both gender-transformative and LNoB approaches were highlighted in the evaluation of phase I of the programme, and efforts have been made to make these a focus in phase II (pp 27). As such this evaluation is an excellent example of one which meaningfully addressed the LNoB agenda throughout this evaluation. The evaluation assess the extent to which gender, marginalised groups, people with disabilities etc, have been served by the programme, in which areas, and where there are areas for improvement. There are multiple examples of where the findings explore how well the intervention has met the needs of the most marginalized, but Finding 3, (pp44-45), explicitly discusses how well the programme reached most at risk groups of girls, for example those living with disabilities, already married girls, LGBTQI people, and those living in different urban or rural settings that contributed to vulnerability, in different country contexts. The humanitarian context is also discussed, and the complexity of reaching girls fleeing e.g. political instability or violence is noted, with reference to the humanitarian development nexus.
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)	9	
	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	GEEW is a core element of this evaluation and is thoroughly integrated into the scope, criteria and questions to ensure that GEEW related data will be collected.
	ii	A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Fully integrated	The methodology is highly gender responsible with robust methods, tools and data analysis approaches that take a gendered approach.
	iii	The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Fully integrated	The findings clearly and systematic apply a gendered analysis. A key point of excellence throughout this report is the way that not just gender, but gender as it intersects with other marginalised factors is explained and analysed. Intersectionality is described and defined (footnote 48 refers to a technical note which in relation to this and the Leave No-one Behind agenda), and as such the evaluation pays close attention, not just to marginalized groups, but to the intersectionality of marginalisation, with gender a key aspect of this for girls who are subject to forced/ child marriage. While multiple findings reference gender, the assessment of the gender transformative approach (summarised in Table 6, pp75,and discussed in detail in relation to finding 13) is a particularly strong element of this evaluation. The conclusions, lessons and recommendations also reflect a gendered analysis, and gender transformative approach.

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?