

GEROS Evaluation Quality Assurance Tool

Version: September 2021

Formative Evaluation of Explosive Ordnance Risk Education programme in Syria

REPORT RATING SUMMARY			
Overall Rating	63%	Satisfactory	
● ● ● ● ●	Exceptional (96% - 100%)	5	
● ● ● ● ●	Highly Satisfactory (87.5% - 95.99%)	4	
● ● ● ● ●	Satisfactory (62.5% - 87.49%)	3	Meets UNICEF/UNEG standards for evaluation reports and decision makers may use the evaluation with confidence
● ● ● ● ●	Fair (35% - 62.49%)	2	
● ● ● ● ●	Unsatisfactory (0% - 34.99%)	1	
REPORT DETAILS			
Title of the evaluation report	Formative Evaluation of Explosive Ordnance Risk Education programme in Syria		
Report sequence number	Syria 59-2024-20129		
Region	MENA		
Year of report	2024		
Office	Syria		
Coverage (countries)	Syria		
ToRs present	Yes		
Date of review (dd/mm/yyyy)	January 14, 2025		
Name of review firm	IOD Parc		
CLASSIFICATION OF EVALUATION REPORT			
Management of evaluation (Managerial control and oversight of evaluation)	UNICEF managed		
Unicef goal areas (Alignment with strategic plan priorities)			
Every child survives and thrives	No		
Every child learns	No		
Every child is protected from violence and exploitation	No		
Every child lives in a safe and clean environment	Yes		
Every child has an equitable chance in life	No		
Gender equality (cross-cutting)	No		
Humanitarian action (cross-cutting)	Yes		
Evaluation object	Programme		
Evaluation type	Formative		
Evaluation strategy	Qualitative		
Evaluation design (primary method used)	Participatory		
Evaluation level	Output & Outcome		
Geographic scope	National		
Primary SDG(s) covered (number)	1, 2, 3, 4, 6, 16		
EQA Summary: The rater will provide top line issues for this evaluation relevant for feedback to senior management (positive and negative), summarizing how the evaluation report meets or fails to meet all criteria. As relevant, the rater will highlight best practice/added value elements and the level of complexity of the evaluation.			
<p>The formative evaluation of the Explosive Risk Ordnance Education (EORE) Programme in the Syrian Arab Republic is an overall 'satisfactory' report with some elements as stronger than others.</p> <p>Strengths of the report include:</p> <ul style="list-style-type: none"> - a strong context section which provides a useful overview of EORE globally and within Syria, and includes SDG targets and indicators as well as the international legal framework. - the report provides a very good sense of the diversity of perspectives, through the analysis and quotes from KIIs and FGDs, which is particularly important given the reliance on primary qualitative data collection. - the process for developing the recommendations was appropriately consultative and iterative, with a likely high degree of ownership. They are well prioritised and targeted. - there is a good linkage between the findings and recommendations, with the findings providing the rationale for the recommendations. <p>Weaknesses include:</p> <ul style="list-style-type: none"> - the methodology could have been more intentional and explicit in terms of how it was going to address gender and other marginalised groups, for example within the evaluation matrix. - the findings section should have adhered to the evaluation criteria and questions throughout (particularly for coherence, efficiency and sustainability). - the report did not put forward a reconstructed ToC, or one for future programming as was required within the ToR, or justify it, but instead proposed it as part of a recommendation. - the findings were uneven in terms of the extent of data and evaluative evidence put forward (particularly around efficiency, coherence and sustainability), and data gaps were not explained and justified. - the conclusions are written and presented as a summary of findings, rather than conclusions and do not provide further strategic insights. - the recommendations do not consistently include practical guidance. 			
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 report could be improved by:</p> <ul style="list-style-type: none"> - the methodology could be revised to be more explicit as to how gender and other marginalised groups (e.g. children with disabilities) would be included, particularly within the evaluation questions. - the evaluation framework could have comprised overarching evaluation questions as well as sub-questions, to help guide the areas of enquiry. - the findings section could have been revisited to adhere to the evaluation matrix list sequence of questions against criteria. - the report could have put forward a reconstructed ToC, and one for future programming as was required within the ToR. - the findings could have included more data and evaluative evidence (particularly around efficiency, coherence and sustainability), and/ or data gaps explained and justified. - the conclusions section could be rewritten to read as conclusions providing additional strategic insights (rather than purely the findings text/ summaries). - the recommendations should all offer practical guidance. 			
SECTION RATINGS			
SECTION A: EXECUTIVE SUMMARY (weight 5%)	100%	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).	Yes	This 4-page executive summary provides a succinct and useful summary of the document, which can be read as a stand alone document to provide the key messages of the report.
	ii	Includes all necessary elements (overview of the intervention, evaluation purpose, objectives and intended audience, evaluation methodology, key conclusions on findings, lessons learned if requested, key recommendations) as per the ToR.	Yes	The summary includes all key elements of the report including overview, evaluation purpose, objectives and audience, methodology, findings, lessons, conclusions and recommendations.
	iii	Includes all significant information needed to understand the intervention and the evaluation AND does not introduce new information from what is presented in the rest of the report.	Yes	The summary provides all information needed to understand the programme, and does not include new information that does not exist in the report.
SECTION B: BACKGROUND (weight 5%)			79%	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 intervention is described in terms of its length (since 2013), overall activities, delivery, coverage and position within UNICEF (paras 23-29). The programme lacks a programme strategy with clearly defined programme objectives, and it defines its outputs, activities in the annual work plans, which are shared. The budget over years and how it has declined is shared in para 28.
	ii	Clear and relevant description of intended rightsholders (beneficiaries) and duty bearers (state and non-state actors with responsibilities regarding the intervention) by type (i.e., institutions/organizations; communities; individuals...), by geographic location(s) (i.e., urban, rural, particular neighbourhoods, town/cities, sub-regions...) and in terms of numbers reached, with disaggregation by gender, age, disability . . . (as appropriate to the purpose of the evaluation).	Partially	Para 27 states the programme coverage in terms of EORE awareness sessions (including for government ministries, implementing partners), capacity development sessions (including for health workers, religious leaders and others) and numbers reached of different beneficiaries, by gender. However, there is no description of these groups as such in terms of who, and where they are and their role/ challenges in EORE. There is no mention of number of people with disabilities reached.
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 thorough description of the issues of explosive weapons, globally and in the specific context of Syria. The risks and harm to children (both direct and indirect), and differential impacts on boys and girls, is explained. There is also a description of relevant policies and institutional arrangements in Syria. The data is referenced using reliable international data sources and data gaps are also highlighted.
	ii	Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Yes	Para 17 and 18 detail the linkages with the SDG and relevant targets and indicators. The text draws upon international research to highlight indirect linkages and all links are mapped in Figure 4.
	iii	Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Yes	The programme focuses on reducing the risk of harm to all children in Syria of explosive weapons, thus using blanket targeting. A description of the beneficiaries as such is therefore not relevant, but the overall plight of children is discussed sufficiently. The report goes as far as possible by considering the subnational levels of risk (para 10), and also the higher proportion of boys than girls that are victims (para 9).
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	Para 19 sets the scene for mine action or Explosive Risk Ordnance Education Programme (EORE) in Syria, setting out the lack of national strategic authority or framework and the fragmented organisations on the ground. In terms of the implementation of the EORE programme, this is broadly described in terms of the field offices and implementing partners across NGOs and implementing partners. However, there is no 'stakeholder map' as such that sets out comprehensively all of the partners.
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Partially	The report describes broadly the key stakeholders (para 25 and 26) and their roles, but does not comprehensively map all of the stakeholders or state their contributions and roles, financial or otherwise.
SECTION C: EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)			67%	Comments on Rating
Question 5.		Is the purpose of the evaluation clearly described?		
	i	Purpose of evaluation is clearly defined, including why it was needed at that point in time, its intended use, and key intended users.	Partially	The evaluation purpose is described. The text varies from that in the ToRs. Although it captures some of the purpose within the ToRs (forward looking evaluation). It also does not articulate why the evaluation is needed at this time i.e. that it comes at an important time to generate evidence and learning to inform the future strategic direction of the UNICEF Syria EORE programme including evidence-based programming and resource mobilisation. The primary users, and their uses, are set out clearly.
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 objectives of the evaluation have been rewritten from the ToR, but it is not made clear as to any changes, why or how they were made. A review of both suggests that the revised objectives are a summary version which contain the key elements, except 'making recommendations for best practices and opportunities for/adjustments needed by UNICEF Syria EORE programme approaches & structures to be the most effective and sustainable in the context of Syria ... and to provide technical support UNICEF Syria EORE programme to create a pathway for implementing those changes including phased objectives towards which UNICEF Syria could design and implement relevant programming and designing the monitoring tools needed'.
	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 shared and adheres to the scope set out in the ToRs in terms of timeframe, thematic area and geographical area.
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 programme's intended results are only set out within the UNICEF Syria CO annual work plan (AWP). There is no programme document setting out the results. Para 29 sets out the relevant annual work plan outputs and indicators.
	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	There is an articulation of the programme logic within para 33. It is an area of enquiry within evaluation questions (EQ) 2 which details behaviour change as the outcome, and outputs being knowledge and skills and critiques the causal relationship.
	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.	No	The evaluation is not a theory-based evaluation (and a theory-based approach was not requested in the ToR). The ToR does however request a (i) reconstructed ToC, and (ii) a readjusted ToC for future programming. Despite this, a reconstructed theory of change and revised ToC are not included within the report, or justified as to why they have not been provided. A new ToC is instead included as part of recommendation 1.
SECTION D:	EVALUATION DESIGN AND METHODOLOGY (weight 20%)	68%	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 relevant criteria are identified and aligned with the questions. The questions relate to both the objectives of evaluability and evaluative assessment, and seems appropriate. Changes have been made to the EQs presented in the ToR, and the process of change articulated which involves appropriate consultation (para 35). However, there is no explanation of the substantial changes within the report which would be expected, particularly give the change in criteria from connectedness (used in humanitarian evaluations) to coherence.
	ii In addition to the questions and sub-questions, the evaluation matrix includes indicators, benchmarks, assumptions and/or other processes from which the analysis can be based and conclusions drawn.	Partially	The evaluation matrix includes questions, but does not include sub questions. It would benefit from over-arching question(s) for each criteria. Indicators are included, although broader benchmarks are not utilised. Data collection methods are included.
Question 9.	Does the report specify adequate methods for data collection, analysis, and sampling?		
	i Evaluation design and set of methods are relevant and adequately robust for the evaluation's purpose, objectives and scope; and are fully and clearly described.	Yes	The evaluation design adheres to the expectations and guidance within the ToR, and appears sufficiently robust for the evaluation purpose, objective and scope.
	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 qualitative only, which is justified that this was the most appropriate given the feasibility of the Syrian context and is in line with the ToR. The number of focus group discussions (FGDs) and key informant interviews (KIs) is sufficient to provide a broad base of perspectives. The data sources are described well. The reviewer would have liked to have read why a survey was not included as a data collection method.

	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 set out clearly for key informant interviews and focus group discussions, in terms of what was planned and achieved (and highlighted as a limitation in para 58). The sex and mean age of participants are recorded. Whilst it is stated that the sampling intended to include children/people living with disabilities, the numbers engaged are not set out.
	iv	Clear and complete description of the methods of analysis.	Partially	The methods of analysis are mentioned but are not fully described, for example para 49 mentions contribution and thematic analysis but does not describe them in detail.
	v	Methodology allows for drawing causal connections between outputs and expected outcomes.	Partially	The expected outputs (from the Syria CO AWP) are set out but the outcomes are not stated, or the links between them. The outcome, and links between outputs and outcome, are stated within a (simple) programme logic in para 33 and which was tested within the evaluation.
	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	The limitations are discussed (paras 57-61) and the implications that these had, as well as how these were addressed.
Question 10.		Are ethical issues and considerations described? The evaluation should be guided by the UNEG ethical standards for evaluation. As such, the evaluation report should include:		
	i	Explicit and contextualized reference to the UNEG obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability) and/or the principles in the 2020 revised UNEG Ethical Guidelines (integrity, accountability, respect, beneficence).	Yes	The appropriate ethical guidance and code of conduct (from UNEG and UNICEF) are referred to. Further, ethical approval was sought and received by the UNICEF ethical review board. Principles are set out and contextualised including do no harm, informed consent, anonymity, safety and security.
	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 set out, including informed consent for adolescents and 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	The evaluation does not include any innovative practice.
SECTION E:		EVALUATION FINDINGS (weight 25%)	67%	Comments on Rating
Question 12.		Do the findings clearly address all evaluation objectives and scope?		
	i	Findings marshal sufficient levels of evidence to systematically address all of the evaluation's questions, sub-questions and criteria.	Partially	The evaluation questions are presented in table 2. Whilst it seems that all of the questions are included and responded to within the findings, there is some re-ordering particularly for efficiency, coherence and sustainability. Notably the question on coherence is now focused upon efficiency. The evaluation question presented in the evaluation matrix is: To what extent has the EORE programme considered and made use of external coordination opportunities and synergies with partners and the wider mine risk sector? How was duplication of activities with other partners or external actors identified and managed and how can coordination help ensure this in the future? However in the main report it is: Q17. To what extent were resources (human, financial) efficiently used to achieve programme results in a timely way, and do tools or plans exist to support efficient implementation? In terms of the marshalling of evidence, the analysis is variable in depth and the degree to which it answers the questions. In some cases (e.g. EQs 3, 4, 5 and 6) whilst the findings are constrained by monitoring data, and lacks other quantitative data (for example, a survey was not undertaken as part of the evaluation), the evaluation uses the available qualitative evidence well to provide diverse perspectives. However, for some questions the responses are very brief and do not provide significant evidence or examples, for example EQ 19 gives a brief overview and does not provide any examples of mechanisms that exist for scalability, and Q20 does not give an examples of cost-effectiveness of scale. Q16 regarding financial allocation does not respond to an element of the question regarding allocation to implementing partners.

	ii	Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Yes	The evaluation tests the causal logic presented in para 33. It is critiqued in EQ2, to inform future programming, and in para 96 the assumptions are analysed and tested.
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	It draws upon the available (qualitative) evidence and triangulates data sources and also highlights any limitations in data. For example para 94 regarding communication materials highlights perspectives from KIIs and FGDs, and also gaps in data (social media analytics) and identifies monitoring improvements that could be made in the future. The lack of quantitative monitoring data is noted for future improvements in programme monitoring. The presentation of FGD soundbites within 1-2 paragraphs, enables demonstration of the diverse perspectives well. However, there are some EQs (e.g. EQ16, EQ19) for which very limited data is put forward.
	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	For some EQs, both strengths and weakness of the programme are identified. The evaluation findings are evidence-based, with a continual narrative around the type and degree of data and evidence. The findings are based around the criteria, questions and indicators outlined in the Evaluation Matrix (Annex 3). However, the evaluation does not present sufficient data for some evaluation questions to respond to them adequately, or highlight data gaps (e.g. EQs 16, 19, 20)
	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 evaluation considers what has been achieved and not, and states why this would be the case. The causal logic is examined (EQ2). However, there are places within the report that could benefit from more thorough analysis of contributory factors, such as para 109 which discusses the decline of coverage, and identifies reduced funding as the key factor. However the analysis does not go beyond this to discuss what has caused the decline in funding (giving an indication of what could be done in the future).
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	Assessment of the programme strategy including the monitoring framework is a key area of enquiry of the evaluation. Sections 6.1 and 6.2 present the findings corresponding to the first objective of the evaluation, delving into the programme's relevance, clarity of logic and results chain, and the adequacy and validity of the indicators, tools, and systems for monitoring, measuring and verifying results.
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 section is mixed in terms of usefulness and formulation. The conclusions are presented in part as general narrative, and as part as summaries of findings (rather than being labelled as conclusions). The structure and formulation is not as expected for a conclusions section, and data sources are referred to rather than relying on more conclusive statements. The degree of depth across the paragraphs varies, and is not reflective of the emphasis that should be placed on different conclusion areas (for example para 154 the EORE materials have limited conclusive discussion). Para 150 is an introduction to the evaluation which is not necessary at this stage of the report.
	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	The conclusions are from the findings, and strengths and limitations are shared. Notably, the text is the same in places as the findings section, and does not provide any fresh or strategic insights.
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 are some concerns about both the content of the 'lessons' and how they are formulated. For example; Paras 164 and 165 provide evaluative assessments about the effectiveness of the EORE programme, rather than being lessons that have applicability beyond the programme. Para 166 provides a reflection on the monitoring system (included elsewhere in the report already) and is not formulated as a lesson. Para 168 has the essence of being a lesson as it has wider applicability but could have been formulated differently, for example ' Collaboration with key Line Ministries is critical for a more integrated approach to EORE and expanding reach and coverage. In this case collaboration with the Ministry of Education and Ministry of Health, has been instrumental in achieving broader reach at the national level and delivering essential life-saving information to children in hard-to-reach areas across sectors. The integration of EORE sessions into the immunisation campaigns led by the MoH and into the school curricula has demonstrated the value of integrated approaches to the delivery of EORE at a greater scale'. Paras 171 and 172 focus more on lessons from the evaluation process, which it would be expected to be included within the methodology section rather than the evaluations.
	ii Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	No	The lessons as currently formulated could not be used by others outside of the programme as principles to bring in to other similar programmes, and do not offer sufficient generalisable information about what has worked well and why.
SECTION G: RECOMMENDATIONS (weight 15%)		63%	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	The recommendations do align with the evaluation purpose, with a focus on both the evaluability and the monitoring framework, and well as the results of the programme itself. They derive logically from the findings, and it is stated as to which findings they draw from.
	ii Recommendations are useful and actionable for primary intended users and uses (relevant to the intervention); guidance is given for implementation, as appropriate.	No	Whilst there are overarching recommendations, only two of eight recommendations (recommendation 2 and recommendation 5) include practical guidance for implementation, and no rationale or explanation is provided as to why. It is critical to provide consistent guidance across the recommendations, for example recommendation 1 could benefit from broad guidance about the scope and process of developing the programme strategy and theory of change.
	iii Process for developing the recommendations is described, and includes the Involvement of duty-bearers, as well as rights holders when feasible (or explanation given for why they were not involved).	Partially	The process for developing the recommendations is clearly explained and shows broader buy-in with appropriate stakeholders. However, the report does not state that rights holders were included, or justify why they were not included.
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 are linked to duty bearers, responsible for each recommendation. They are prioritised in terms of importance and timeframe.
SECTION H: REPORT STRUCTURE AND PRESENTATION (weight 5%)		79%	Comments on Rating
Question 19. Does the evaluation report include all relevant information?			
	i Opening pages include: Name of evaluated object, timeframe of the evaluation, date of report, location of evaluated object, names and/or organization(s) of the evaluator(s), name of organization commissioning the evaluation, table of contents -including, as relevant, tables, graphs, figures, annexes; list of acronyms/abbreviations, page numbers.	Yes	The opening pages include the name of the programme, the date of the report, the location, name of the evaluator, table of contents, list of acronyms. It also helpfully includes a glossary table.
	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.	Partially	Annexes include a terms of reference, list of interviewees, evaluation matrix, list of FGDs. However it does not include a list of documentary evidence (beyond the end notes). It includes an annex with the data collection instruments.
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 report is easy to navigate , with clear titles and subtitles and numbered paragraphs.

	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 structure aligns to UNICEF guidelines, and follows the format of content, purpose and methodology, findings, conclusions and recommendations. However there are notable aspects within two of the sections. The conclusions section is presented as a summary of key findings. Also within the findings section, there is some re-ordering of questions which leads to some mismatching of criteria and questions.
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 main body of the report (excluding executive summary and annexe) is 48 pages which is considered a reasonable length (and there is no specification in the ToR).
	ii Report is easy to understand (written in accessible way for intended audience) and generally free from grammar, spelling and punctuation errors.	Partially	The report is generally well written (particularly the context section) and accessible. There are a few typos/ errors. Paragraph 56 has 'protection' at the start, which belongs in paragraph 55. Para 63 has text 'see Error! Reference source not found' as do some of the recommendations (1, 2 and 4).
	iii Frequent use of visual aids (such as infographics, maps, tables, figures, photos) to convey key information. These are clearly presented, labelled, and referenced in text.	Yes	The text is broken up well with tables, graphs and some extracts from FGDs, which are clearly presented, labelled and referenced in text.
SECTION I:	EVALUATION PRINCIPLES (weight 10%)	50%	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 references the CRC as well as other international treaties related to explosive weapons, such as the Convention on Cluster Munitions (CCM) or the Convention on Certain Conventional Weapons (CCW). Pillars of action in mines are referred to as an international framework (paras 14-16).
	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	There are statements that the evaluation reference group (ERG) was consulted for the evaluation questions and recommendations. However, there is no information of who the ERG comprises. There is a description of how right holders were included within the data collection process with FGDs (disaggregated by sex).
	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	The language is inclusive and does not exhibit any bias. As above, rights holders were included in in data collection processes and data disaggregated by sex, however the language of right holders and duty bearers is not used.
	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	Whilst there are attempts to include marginalised groups, this was challenging given the lack of disaggregated data. In terms of data collection, numbers are recorded by sex, (and mean age) but the report does not provide any data about other groups such as children with disabilities. A disability lens is utilised with analysis of the presence of a disability strategy and different needs.
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)	3	
	i GEEW is integrated in the Evaluation Scope of analysis, and evaluation criteria and questions are designed in a way that ensures GEEW-related data will be collected.	Partially integrated	The evaluation matrix includes disaggregation by subgroups/ vulnerable groups for some questions but does not explicitly state that it should be on gender (and other) grounds.
	ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Partially integrated	The methodology does not state its intention as regards integrating gender within the data collection processes and analysis. The number of evaluation participants is disaggregated by male, female, and most FGDs were held as mixed gender (only one governorate were they held as single sex FGDs). The limited participation of boys and males within the evaluation (particularly given that they are disproportionately affected) is noted as a limitation of the evaluation.
	iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Partially integrated	A gender lens is used within the analysis, exploring the differential impacts for boys and girls. The findings is that boys are more directly affected by explosive risk, and the indirect effects for girls and boys are also considered. Whilst the data suggests that participation by boys and girls in the programme is equal, the limited participation of boys in the evaluation process meant that it was difficult to verify this. The conclusions and recommendations do not discuss any gender related findings.

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?