

GEROS Evaluation Quality Assurance Tool

Version: November, 2020

Evaluation of UNICEF's Makani Programme in Jordan (January 2018-January 2019)

REPORT RATING SUMMARY			
Overall Rating		91%	Highly Satisfactory
●●●●●	Exceptional (96% - 100%)	5	
●●●●●	Highly Satisfactory (87.5% - 95.99%)	4	Exceeds UNICEF/UNEG standards for evaluation reports and decision makers may use the evaluation with a high degree of confidence
●●●●-	Satisfactory (62.5% - 87.49%)	3	
●●●-●	Fair (35% - 62.49%)	2	
●●-●●	Unsatisfactory (0% - 34.99%)	1	
REPORT DETAILS			
Title of the evaluation report		Evaluation of UNICEF's Makani Programme in Jordan (January 2018-January 2019)	
Report sequence number		Jordan/59/2020/17085	
Region		MENA	
Year of report		2020	
Office		UNICEF Jordan	
Coverage (countries)		Jordan	
ToRs present		Yes	
Date of review (dd/mmm/yyyy)		January 12, 2021	
Name of review firm		DefEdge	
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		Yes	
Every child is protected from violence and exploitation		Yes	
Every child lives in a safe and clean environment		Yes	
Every child has an equitable chance in life		Yes	
Gender equality (cross-cutting)		Yes	
Humanitarian action (cross-cutting)		Yes	
Evaluation object		Programme	
Evaluation type		Formative	
Evaluation strategy		Mixed methods	
Evaluation design (primary method used)		Theory-based	
Evaluation level		Output & Outcome	
Geographic scope		National	
Primary SDG(s) covered (number)		Not specified	
EQA Summary:			
<p>This is a strong evaluation of the Makani programme which is implemented in the context of the protracted conflict and humanitarian crisis in Syria and consequent influx of refugees to Jordan. It seeks to assess the extent to which the intervention helped vulnerable children and youth in achieving their full potential through participating in the Makani interventions. This formative evaluation intends to improve Makani's future interventions by re-programming processes through identification and correction of ineffective practices. The evaluation uses a solid participatory and mixed-method approach involving a large sample of implementors and rights holders. To assess quantitative data, the evaluation uses exploratory, descriptive analysis and a random effects (RE) regression analysis. It also thoroughly addresses gender and vulnerable groups. The report comprehensively covers all important aspects, however it is significantly exceeds the page length requested in the ToR.</p>			
Recommendations for Improvement:			
<p>In terms of presentation, the main issue is the lengthiness of the report. Both the context and findings sections were quite detailed, and some of this information could have been written more concisely with parts relegated to the annex. It is good practice to make explicit and contextualized reference to the obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability). The lessons provide useful insights learnt from the evaluation; however, they could have been framed more broadly to have applicability to other interventions.</p>			
SECTION RATINGS			
SECTION A:	EXECUTIVE SUMMARY (weight 5%)	83%	Comments on Rating (include explanations for any criterion not rated)
Question 1.	Can the executive summary inform decision-making?		
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	The executive summary length is 3.5 pages.
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	All required elements are included except for the intended audience. Lessons learned is also not present in the summary however it is not required by the ToR
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 evaluation context, purpose, methodology and results are clear and the report does not introduce new information from the rest of the report.
SECTION B:	BACKGROUND (weight 5%)	86%	Comments on Rating (include explanations for any criterion not rated)
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 background, context, and evolution of Makani programme clearly described. The budget and UNICEF contribution is also given.

	ii Clear and relevant description of intended rightsholders (beneficiaries) and duty bearers (state and non-state actors with responsibilities regarding the intervention) by type (i.e., institutions/organizations; communities; individuals...), by geographic location(s) (i.e., urban, rural, particular neighbourhoods, town/cities, sub-regions...) and in terms of numbers reached, with disaggregation by gender, age, disability... (as appropriate to the purpose of the evaluation).	Yes	There is a clear description given of both rightsholders and duty bearers (including other actors involved in different capacities). Rightsholders are categorized by vulnerability. For example: (1) children not enrolled in schools, (2) adolescent girls, (3) children unaccompanied by their families or separated from them, (4) children with special needs, (5) children survivors of gender-based violence, including early marriage, (6) children who are begging, working or at risk of begging or working, (7) children of extremely poor families, (8) children of female-headed households, and (9) boys and girls at risk of the impact of extreme views.
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	The context, especially in terms of humanitarian crisis due to refugee influx from Syria, is well described.
	ii Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	No	The intervention's contributions to SDGs in general are assessed, however the relevant SDGs are not actually specified.
	iii Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Yes	All beneficiaries are defined and categorized.
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	All implementing partners and other actors involved in implementation of intervention are carefully defined.
	ii Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Yes	The roles of government, Ministry of Social Development (MoSD) and five other NGOs (IPs), and other stakeholders (incl. an international NGO and a research institution) are clear.
SECTION C:	EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)	100%	Comments on Rating (include explanations for any criterion not rated)
Question 5.	Is the purpose of the evaluation clearly described?		
	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 and use of the evaluation, as well as primary and secondary intended users are clearly defined.
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 complete description of the objectives and intended use of this formative evaluation. This evaluation intends to support the Makani intervention's future programming processes by identifying ineffective practices for correction. Moreover, the results of this evaluation are intended to guide the UNICEF Jordan Country Programme to reposition and strategic shift of the perspective for the rest of the programme cycle.
	ii Clear and relevant description of the scope of the evaluation: what will and will not be covered (thematically, chronologically, geographically with key terms defined), as well as, if applicable, the reasons for this scope (e.g., specifications by the ToRs, lack of access to particular geographic areas for political or safety reasons at the time of the evaluation, lack of data/evidence on particular elements of the intervention).	Yes	The scope is in alignment with the ToR
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	There is a clear description of intervention's intended result is provided in theory of change. It is shown in Figure 2
	ii Causal relationship between outputs and outcomes is presented in narrative and/or graphic form (e.g., results chain, logic model, theory of change, evaluation matrix).	Yes	The causal relationship between outputs and outcomes is presented in the theory of change model.
	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 evaluation verified, updated, and validated the theory of change with key stakeholders. The evaluation team also identified and assessed context-dependent and exogenous effects that influenced ToC using a Context, Mechanism, Outcome (CMO) model.
SECTION D:	EVALUATION DESIGN AND METHODOLOGY (weight 20%)	90%	Comments on Rating (include explanations for any criterion not rated)
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.	Yes	The questions follow the OECD/DAC criteria, except coherence. The ToR did not require 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.	Yes	Annex 1 includes a detailed evaluation matrix with indicators, data sources, collection methods, and comments on 'What we are looking at • What we are looking for • Key areas of enquiry'.
Question 9.		Does the report specify adequate methods for data collection, analysis, and sampling?		
	i	Evaluation design and set of methods is relevant and adequately robust for the evaluation's purpose, objectives and scope; and are fully and clearly described.	Yes	The evaluation design uses a standard mixed-method approach with both qualitative and quantitative methods. Data collection was extensive and well-described. Since the evaluation involved key informant interviews and around 100 focus groups with target children, it is primarily qualitative.
	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 include beneficiaries and implementors and they are clearly described by type, location and gender. Sources of documentary review are also provided. Quantitative data is gathered from the Bayanati database.
	iii	Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or if not, provide reasons for this).	Yes	The evaluation uses a stratified sampling strategy. For national level KIs, sampling is based on 'function', 'need', as well as 'gender balance'. The report mentions that sampling of participants is based on age, gender, nationality, vulnerabilities and types of services accessed or provided. Importantly, specific FGDs were conducted with children belonging to the Dom minority and children who dropped out of formal education.
	iv	Clear and complete description of the methods of analysis.	Yes	The methods of analysis are carefully described and included random effects regression analysis for qualitative data.
	v	Methodology allows for drawing causal connections between outputs and expected outcomes.	Yes	The connections are clear.
	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	Limitations as well as mitigation strategies are carefully described, including in reference to evaluation timeframe, availability of quantitative data, and implementation of an integrated approach.
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 obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability).	No	These are not provided.
	ii	Description of ethical safeguards for participants appropriate for the issues relevant to methodology and how they are applied (respect for dignity and diversity, right to self-determination, fair representation, compliance with codes for vulnerable groups, confidentiality, and avoidance of harm). For those cases where the evaluation involves interviewing children, explicit reference is made to the UNICEF procedures for Ethical Research Involving Children.	Yes	There is a detailed description given on how the ethical standards were applied including receipt of ethical approval in conformance with ERIC.
SECTION E: EVALUATION FINDINGS (weight 25%)			92%	Comments on Rating (include explanations for any criterion not rated)
Question 11.		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 findings include substantial evidence. Qualitative evidence such as quotes are provided in boxes. However, the presentation would be more clear if the evidence was presented against each evaluation question.
	ii	Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Yes	The ToC was analysed and validated with the key stakeholders. In addition, effectiveness section provides findings on output and outcome targets as stated in the ToC (QE3).
Question 12.		Are evaluation findings derived from the conscientious, explicit and judicious use of the best available, objective, reliable and valid data and by accurate quantitative and qualitative analysis of evidence.		
	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	Evaluation findings are derived from interviews, focus groups, as well as quantitative data. Triangulation can be observed throughout.
	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 use both the 2018 results matrix indicators as well as the indicators from 2018 to derive positive as well as negative findings.
	iii	Causal factors (contextual, organizational, managerial, etc.) leading to achievement or non-achievement of results are clearly identified. For theory-based evaluations, findings analyse the logical chain (progression -or not- from implementation to results).	Yes	Causal connections are specified in the theory of change. Findings describe reasons behind achievement and non-achievement of results.
Question 13.		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 if required in ToR) to support decision-making.	Yes	There is a thorough assessment of M&E mechanisms and systems.

SECTION F:	EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 10%)	75%	Comments on Rating (include explanations for any criterion not rated)
Question 14.	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	Conclusions are clear, purposeful, and forward looking. For example, "The programme's highly developed monitoring and evaluation arrangement offers a wealth of frequently updated information which can feed into the decision-making process".
	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 from findings. There are separate sections on strengths and weaknesses.
Question 15.	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.	Partially	The lessons provide useful insights learnt from the evaluation, however the lessons could have been reframed for wider applicability. For example, it is noted that there is potential and rationale for Makani transposition in other humanitarian and development contexts however it is not clear how and why it is relevant.
	ii Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Yes	Lessons are clear, concise, and are useful for the intended audience to improve programming
SECTION G:	RECOMMENDATIONS (weight 15%)	100%	Comments on Rating (include explanations for any criterion not rated)
Question 16.	Are recommendations well grounded in the evaluation?		
	i Recommendations align with the evaluation purpose, are clearly formulated and logically derived from the findings and/or conclusions.	Yes	Recommendations are derived from findings and are well presented.
	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.
	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).	Yes	The participatory approach is carefully described and is demonstrative of good practice. Views of beneficiaries collected during fieldwork were incorporated. In addition, a one-day workshop devoted to the discussion of the preliminary findings and development of operationally feasible recommendations was organised.
Question 17.	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	Recommendations are prioritized and directed towards responsible entities.
SECTION H:	REPORT STRUCTURE AND PRESENTATION (weight 5%)	86%	Comments on Rating (include explanations for any criterion not rated)
Question 18.	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 required information appears.
	ii Annexes include, if not in report body: terms of reference, evaluation matrix, list of interviewees, results chain/ToC/logical framework, 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, results chain, information about the evaluator(s).	Yes	The annexes include all required elements.
Question 19.	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.
	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 usual structure is followed.
Question 20.	Is the report well presented?		
	i Report is of reasonable length; it does not exceed number of pages that may be specified in ToR.	No	The ToR required the report to be between 40-60 pages. However at 86 pages, it is overly lengthy.
	ii Report is easy to understand (written in accessible way for intended audience) and generally free from grammar, spelling and punctuation errors.	Yes	It is easy to understand and free of errors
	iii Frequent use of visual aids (such as infographics, maps, tables, figures, photos) to convey key information. These are clearly presented, labeled, and referenced in text.	Yes	There was a good range of visual aids used including photographs, conceptual diagrams and charts.
SECTION I:	EVALUATION PRINCIPLES (weight 10%)	95%	Comments on Rating (include explanations for any criterion not rated)
Question 21.	Did the evaluation design and style consider incorporation of the UN and UNICEF's commitment to a human rights-based approach to programming, to gender equality, and to equity?		
	i Reference and use of rights-based framework, and/or CRC, and/or CCC, and/or CEDAW and/or other rights related benchmarks in the design of the evaluation.	Yes	References are made to human rights frameworks including the CRC.

	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).	Yes	The evaluation followed a participatory approach and engaged stakeholders and beneficiaries in the research process and in the formulation of recommendations. A stakeholder workshop was also conducted at the end to validate findings and recommendations.
	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	The report is written in gender-sensitive and human-rights sensitive manner. Data is disaggregated by age, gender, and nationality. The evaluation has also assessed issues children face due to their background. For example, the report mentions that "Once established in a new location, Dom parents are often told their children are ineligible for schooling". In addition, the report also notes that the majority of parents who benefited from the programme were not engaged in labour (61%) and were women (82% females), highlighting cultural role as well as capacity issues.
	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 evaluation describes issues of different marginalized groups and assess the extent to which these groups were reached. For example the report mentions that "Dom minority face not only severe levels of poverty (resulting in prevalence of child labour and early marriages), but also deal with social and cultural stigma, with the lack of sanitation and hygiene facilities". Issues of disability are also addressed.
Question 22.	Does the evaluation meet UN SWAP evaluation performance indicators? (Note: this question will be rated according to UN SWAP standards)	8	
	i GEEW is integrated in the Evaluation Scope of analysis, and evaluation criteria and questions are designed in a way that ensures GEEW-related data will be collected.	Fully integrated	The objective included the need to assess the extent to which the programme integrated gender, equity and child rights. GEEW is mainstreamed into Relevance and Effectiveness, and addressed under x-cutting issues. There is an emphasis on collecting gender-disaggregated data.
	ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Satisfactorily integrated	The evaluators state that a child-sensitive and gender-specific approach was applied. The approach was participatory and mixed methods; ethical practices were clearly described, however there were none explicitly gender focussed. While the data was disaggregated by gender in many cases, including the number of FGD participants, the total number of evaluation participants was not.
	iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Fully integrated	The background section discusses the situation of girls/women and issues faced. A gender analysis is reflected in the findings (for example, that there was no real difference found in improvement in mathematical skills between boys and girls), conclusions, and recommendations.
SWAP Rating Guidance			
i GEEW is integrated in the Evaluation Scope of analysis, and evaluation criteria and questions are designed in a way that ensures GEEW-related data will be collected.			
a. Does the evaluation assess whether sufficient information was collected during the implementation period on specific result indicators to measure progress on human rights and gender equality results?			
b. Does the evaluation include an objective specific to assessment of human rights and gender equality considerations or was it mainstreamed in other objectives?			
c. Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria?			
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?			