

GEROS Evaluation Quality Assurance Tool

Version: February, 2020

Developmental Evaluation of Unicef's Work on Innovation for Children in Malaysia

REPORT RATING SUMMARY			
Overall Rating		91%	Highly Satisfactory
● ● ● ●	Highly Satisfactory (87.5% - 100%)	4	Exceeds UNICEF/UNEG standards for evaluation reports and decision makers may use the evaluation with a high degree of confidence
● ● ● ●	Satisfactory (62.5% - 87.49%)	3	
● ● ● ●	Fair (34.76% - 62.49%)	2	
● ● ● ●	Unsatisfactory (0% - 34.75%)	1	
REPORT DETAILS			
Title of the evaluation report		Developmental Evaluation of Unicef's Work on Innovation for Children in Malaysia	
Report sequence number		Malaysia/60/2019/15039	
Region		EAPR	
Year of report		2019	
Office		Malaysia Country Office	
Coverage (countries)		Malaysia	
ToRs present		Yes	
Date of review (dd/mmm/yyyy)		March 16, 2020	
Name of review firm		AIMS	
CLASSIFICATION OF EVALUATION REPORT			
Management of evaluation (Managerial control and oversight of evaluation decisions)		UNICEF managed	
Unicef goal areas			
Every child survives and thrives		Yes	
Every child learns		Yes	
Every child is protected from violence and exploitation		Yes	
Every child lives in a safe and clean environment		Yes	
Every child has an equitable chance in life		Yes	
Gender equality (cross-cutting)		Yes	
Humanitarian action (cross-cutting)		No	
Evaluation object		Country Programme	
Evaluation type		Formative	
Evaluation strategy		Mixed methods	
Evaluation design (primary method used)		Participatory	
Evaluation level		Output	
Geographic scope		National	
Primary SDG(s) covered (number)		general applicability to multiple SDGs	
EQA Summary: <i>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.</i>			
<p>This is a strong developmental evaluation aimed at identifying new opportunities to use innovation to increase the impact of the Malaysia Country Programme. The methodology was highly participatory and included the use of both external and internal embedded evaluators to help drive what was primarily a learning process. It was suitably robust for purpose and included document review, sourcing case studies and examples of innovation practice, KIs, observation, and an innovation workshop. A learning framework, developed by the consultant company, guided the evaluation/learning process. The ToC developed by the evaluators and 33 findings are clearly and succinctly presented with good use of graphics. The conclusions and recommendations appear very useful for moving the innovation agenda ahead; particularly as clear operational suggestions are provided for each recommendation (and are an example good practice). The Executive Summary could have been more clear by providing a summary of findings rather than listing all 33, and by providing concrete examples of innovation practice that would be understood by internal and external audiences. The evaluators could also have brought a stronger gender lens to the study - although the methodology was gender responsive, there was minimal incorporation of GEEW in the learning framework and findings, and no reflection of gender in the theory change developed or the conclusions and recommendations.</p>			
Recommendations for Improvement: <i>The rater will identify topline recommendations that would improve the evaluation, and be specific to the sections of the report where shortcomings were found. Cite resources that may assist evaluation managers in overseeing future evaluations as relevant.</i>			
<p>There are two suggested areas for improvement. It would have been helpful for the Executive Summary to have better defined innovation, especially as a key finding was that the concept of innovation remains unclear to many stakeholders. The summary could have referenced some of the innovative practices provided later in the report to better set the context and to help ensure that all those who just read the summary would understand the object of the evaluation. External audiences would not necessarily be familiar with the one example cited, U-reports. The evaluation's analysis of gender was quite light; it looked at whether HRGE was considered in the design of innovation tools, but did not go further to suggest how this could be done or investigate whether/how innovation may be taken up and experienced differently by male and female stakeholders.</p>			
SECTION RATINGS			
SECTION A:	EXECUTIVE SUMMARY (weight 5%)	67%	Comments on Rating (include explanations for any criterion not rated)
Question 1.	Can the executive summary inform decision-making?	67%	
	An executive summary is provided that is of relevant conciseness and depth for primary intended users (Maximum of approximately 5 pages unless otherwise specified in ToR).	Partially	The presentation is 5 pages. Although the depth is appropriate, the main findings are presented in 33 bullet points, and could have been more clear if synthesized.

	ii	Includes all necessary elements (overview of the intervention, evaluation purpose, objectives and intended audience, evaluation methodology, key findings, key conclusions, key recommendations) as per the ToR.	Yes	All of the elements are covered, including an explanation of Development Evaluation.
	iii	Includes all the necessary information to understand the intervention and the evaluation AND does not contain information not already included in the rest of the report.	Partially	This section could more clearly communicated what 'innovation' is by giving examples of good practice. One example, U-Report, is cited but is not explained, limiting its usefulness for identified secondary users of the report who are outside the UN system. The evaluation process is covered but the EE position is not explained and is only referred to by the acronym.
SECTION B: BACKGROUND (weight 5%)			92%	Comments on Rating (include explanations for any criterion not rated)
Question 2. Is the object of the evaluation clearly described?			100%	
	i	Clear and relevant description of the intervention, including: location(s), timelines, cost/budget, and implementation status.	Yes	There is an explanation of innovation within UNICEF as a whole and within UNICEF Malaysia. This includes timeframe and implementation status. Financial information is presented later in Findings.
	ii	Clear and relevant description of intended beneficiaries by type (i.e., institutions/organizations; communities; individuals...), by geographic location(s) (i.e., urban, rural, particular neighbourhoods, town/cites, sub-regions...) and in terms of numbers reached (as appropriate to the purpose of the evaluation).	Yes	Although not explicitly stated in this section, it is clear that there is perceived potential benefit for all UNICEF rights holders and duty bearers.
Question 3. Is the context of the intervention clearly described?			100%	
	i	Clear and relevant description of the context of the intervention (relevant policy, socio-economic, political, institutional, international factors) and how context relates to the implementation of the intervention.	Yes	The context is shown as the UNICEF-Malaysia Country Plan; its alignment with the Government of Malaysia plan and key pillars is provided. The high-level results of the recent global UNICEF evaluation on innovation provide additional context.
	ii	Clear and relevant description (where appropriate) of the status and needs of the target groups for the intervention.	Yes	There is an adequate overview of the four main groups of intended rights holders of Unicef's work in the area of innovation: refugees, migrants and stateless persons; urban poor; adolescents; and children with disabilities.
Question 4. Are key stakeholders and their contributions clearly identified?			75%	
	i	Identification of implementing agency(ies), development partners, primary duty bearers, secondary duty bearers, and rights holders.	Yes	In describing the global-level evaluation, potential innovative types of partnerships are given (start ups, private sector, academics, international agencies). Examples of key stakeholders in Malaysia are sprinkled throughout the section.
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Partially	The roles of UNICEF's Innovation Specialist is given. Some examples are given of innovative initiatives supported by UNICEF Malaysia but specific roles, including of other implementors are not noted. Contributions and roles are more clear in Boxes 3 - 6 which provide examples of innovation in other countries.
SECTION C: EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)			100%	Comments on Rating (include explanations for any criterion not rated)
Question 5. Is the purpose of the evaluation clearly described?			100%	
	i	Specific identification of how the evaluation is intended to be used and to what this use is expected to achieve.	Yes	The purpose is identified as helping UNICEF identify new opportunities to increase impact of the CP through innovation. Uses are also specified.
	ii	Identification of appropriate primary intended users of the evaluation.	Yes	Primary and secondary users are identified.
Question 6. Are the objectives and scope of the evaluation clear and realistic?			100%	
	i	Clear and complete description of what the evaluation seeks to achieve by the end of the process with reference to any changes made to the objectives included in the ToR (if applicable).	Yes	The three objectives of the development evaluation are given.
	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 specified in terms of timeframe, location of activities considered, and forward-looking aim.
Question 7. Is the results chain or logic well articulated?			100%	
	i	Clear description of the intervention's intended results, or of the parts of the results chain that are applicable to the evaluation.	Yes	This is done in terms of the four critical assumptions underpinning the ToC.
	ii	The causal relationship between outputs and outcomes is presented in narrative and/or graphic form (theory of change, logic model, results chain, evaluation matrix, etc.).	Yes	There is a clear narrative description and diagram of the ToC developed by the embedded evaluators (EEs).

SECTION D: EVALUATION METHODOLOGY (weight 20%)		93%	Comments on Rating (include explanations for any criterion not rated)
Question 8.	Does the evaluation provide a relevant list of evaluation criteria that are explicitly justified as appropriate for the purpose of the evaluation? UNICEF evaluation standards refer to the OECD/DAC criteria. Not all OECD/DAC criteria are relevant to all evaluation objectives and scopes. Standard OECD DAC Criteria include: Relevance; Effectiveness; Efficiency; Sustainability; Impact. Evaluations should also consider equity, gender and human rights (these can be mainstreamed into other criteria). Humanitarian evaluations should also consider Coverage; Connectedness; Coordination; Protection; Security.	100%	
i	Clear and relevant presentation of the evaluation framework including clear evaluation questions used to guide the evaluation.	Yes	The Dublin Framework for Innovation is noted as being the learning framework used. Its 4 dimensions and corresponding areas of enquiry and questions are given. Box 7 provided a useful comparison between developmental and traditional evaluations.
Question 9.	Does the report specify methods for data collection, analysis, and sampling?	100%	
i	Clear and complete description of a relevant design and set of methods that are suitable for the evaluation's purpose, objectives and scope.	Yes	The design and the different phases of the evaluation are clearly explained and suitable.
ii	Clear and complete description of the data sources, rationale for their selection and sampling strategy. This should include a description of how diverse perspectives are captured (or if not, provide reasons for this), how accuracy is ensured, and the extent to which data limitations are mitigated.	Yes	Data sources included desk review, searches for innovative practice, KIIs, observations, and a workshop to further innovation, all leading to a current state assessment of UNICEF's work on innovation. The purposive sampling strategy was described. Interestingly, although there is a very good graphic depiction of types of stakeholder respondents broken down by percent, the total number of respondents is not given in this section (although the number is provided in the Executive Summary and can later be found buried in the annex).
iii	Clear and complete description of the methods of analysis, including triangulation of multiple lines and levels of evidence (if relevant).	Yes	The analysis methods included coding of KII responses and documents. The triangulation process is explained.
iv	Clear and complete description of limitations and constraints faced by the evaluation, including gaps in the evidence that was generated and mitigation of bias, and how these were addressed by the evaluators (as feasible).	Yes	Limitations and mitigation actions are adequately explained.
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:	75%	
i	Explicit reference to the obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability).	Yes	UNEG ethical standards are covered and include how each obligation was enacted.
ii	Description of ethical safeguards for participants appropriate for the issues described (respect for dignity and diversity, right to self-determination, fair representation, compliance with codes for vulnerable groups, confidentiality, and avoidance of harm). For those cases where the evaluation involves interviewing children, explicit reference is made to the UNICEF procedures for Ethical Research Involving Children.	Partially	It is less clear how ethical safeguards were upheld, but reference is made to relevant guidance.
SECTION E: EVALUATION FINDINGS (weight 25%)		93%	Comments on Rating (include explanations for any criterion not rated)
Question 11.	Do the findings clearly address all evaluation objectives and scope?	100%	
i	Findings marshal sufficient levels of evidence to systematically address all of the evaluation's questions and criteria.	Yes	All questions are systematically addressed.
ii	Reference to the intervention's results framework in the formulation of the findings.	Yes	The learning framework is used.
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.	100%	
i	The evaluation clearly presents multiple lines (including multiple time series) and levels (output, outcome, and appropriate disaggregation) of credible evidence.	Yes	The findings draw on evidence from the different data sources. Findings are mainly qualitative in nature but there is some use of quantitative data. Document citations are given.
ii	Findings are clearly supported by and respond to the evidence presented, including both positive and negative. Findings are based on clear performance indicators, standards, benchmarks, or other means of comparison.	Yes	Findings respond to both positive and negative evidence. Findings are generally based on the assumptions (benchmarks) posed in the learning framework which form the basis of the questions.
iii	The causal factors (contextual, organizational, managerial, etc.) leading to achievement or non-achievement of results are clearly identified. For theory-based evaluations, findings analyse the logical chain (progression -or not- from implementation to results).	Yes	Causal factors, such as organizational processes, are clearly brought out.

Question 13.	Does the evaluation assess and use the intervention's Results Based Management elements?	75%	
	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).	Partially	This is indirectly assessed. One question asks whether there are defined metrics around innovation; the finding notes that lack of metrics creates a disincentive for innovation.
	ii Assessment of the use of monitoring data in decision making.	Yes	As this is a formative evaluation, assessment of use in decision making is not yet relevant.
SECTION F:	EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 10%)	100%	Comments on Rating (include explanations for any criterion not rated)
Question 14.	Do the conclusions present an objective overall assessment of the intervention?	100%	
	i Clear and complete description of the strengths and weaknesses of the intervention that adds insight and analysis beyond the findings.	Yes	The conclusions are clearly presented and provide a good summary of constraints and opportunities.
	ii Description of the foreseeable implications of the findings for the future of the intervention (if formative evaluation or if the implementation is expected to continue or have additional phase).	Yes	The conclusions are all forward looking, in line with expectations of a developmental evaluation.
	iii The conclusions are derived appropriately from findings.	Yes	They clearly flow from findings.
Question 15.	Are logical and informative lessons learned identified?	100%	
	i Identified lessons that stem logically from the findings, presents an analysis of how they can be applied to different contexts and/or different sectors, and takes into account evidential limitations such as generalizing from single point observations.	Yes	The lessons are informative and well framed. In addition to being useful for other situations, this section provides brief examples of how other UN agencies promote innovation.
SECTION G:	RECOMMENDATIONS (weight 15%)	90%	Comments on Rating (include explanations for any criterion not rated)
Question 16.	Are recommendations well grounded in the evaluation?	83%	
	i Recommendations are logically derived from the findings and/or conclusions.	Yes	The recommendations flow from findings and conclusions.
	ii Recommendations are useful to primary intended users and uses (relevant to the intervention and provide realistic description of how they can be made operational in the context of the evaluation).	Yes	These appear useful. The highlighting of operational suggestions for each is an example of best practice.
	iii Clear description of the process for developing recommendations, including a relevant explanation if the level of participation of stakeholders at this stage is not in proportion with the level of participation in the intervention and/or in the conduct of the evaluation.	Partially	It is noted in methodology that the Reference Group was involved in several rounds of review of the report but their involvement in the development of recommendations is not explicitly addressed.
Question 17.	Are recommendations clearly presented?	100%	
	i Clear identification of target group for action for each recommendation (or clearly clustered group of recommendations).	Yes	Target groups are clearly specified.
	ii Clear prioritization and/or classification of recommendations to support use.	Yes	High and medium priority ratings are given.
SECTION H:	EVALUATION STRUCTURE/PRESENTATION (weight 5%)	80%	Comments on Rating (include explanations for any criterion not rated)
Question 18.	Does the evaluation report include all relevant information?	75%	
	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	All information is provided.
	ii Annexes should include, when not present in the body of the report: Terms of Reference, Evaluation matrix, list of interviewees, list of site visits, data collection instruments (such as survey or interview questionnaires), list of documentary evidence. Other appropriate annexes could include: additional details on methodology, copy of the results chain, information about the evaluator(s).	Partially	The only missing required element appears to be the interviews protocols for KIIs. The presentation of the annexes is not very clear. They are listed in the table of contents by evaluation phase rather than item, making it difficult to find, for example, the list of KII respondents.
Question 19.	Is the report logically structured?	100%	
	i The structure is easy to identify and navigate (for instance, with numbered sections, clear titles and sub-titles).	Yes	The report is clearly structured.
	ii Context, purpose and methodology would normally precede findings, which would normally be followed by conclusions, lessons learned and recommendations.	Yes	The regular order is followed.
Question 20.	Is the report well presented?	50%	
	i 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 clearly written with few errors. Strong graphics bring clarity. However, narrow page margins and the 2-column format make some sections quite text heavy. A further issue is that the report exceeds the 40 page requested limit in the ToR by over 20 pages.
SECTION I:	EVALUATION PRINCIPLES (weight 10%)	83%	Comments on Rating (include explanations for any criterion not rated)
Question 21.	Did the evaluation design and style consider incorporation of the UN and UNICEF's commitment to a human rights-based approach to programming, to gender equality, and to equity?	100%	

	i Reference and use of rights-based framework, and/or CRC, and/or CCC, and/or CEDAW and/or other rights related benchmarks in the design of the evaluation.	Yes	The CRC, CCC an CEDAW are cited as informing the methodology. The scope of the assignment is explicit about informing ways innovation could add value to Unicef's rights-based and other work.
	ii Clear description of the level of participation of key stakeholders in the conduct of the evaluation, (for example, a reference group is established, stakeholders are involved as informants or in data gathering).	Yes	UNICEF stakeholders were closely involved in this participatory process, including as EEs and Reference Group members,
	iii Stylistic evidence of the inclusion of these considerations which can include: using human-rights language; gender-sensitive and child-sensitive writing; disaggregating data by gender, age and disability groups; disaggregating data by socially excluded groups.	Yes	Rights-based language is used throughout. The generally report reflects UNICEF's commitment to HRBA.
Question 22.	Does the evaluation assess the extent to which the implementation of the intervention addressed gender, equity & child rights?	#DIV/0!	
	i Identification and assessment of the presence or absence of equity considerations in the design and implementation of the intervention (such as the involvement in the intervention of right holders, duty bearers, and socially marginalized groups, and the differential benefits received by different groups of children).	Not Rated	As the evaluation is looking at internal processes, this question is considered not relevant.
	ii Clear proportionality between the level of participation in the intervention and in the evaluation, or clear explanation of deviation from this principle (this may be related to specifications of the ToRs, inaccessibility of stakeholders at the time of the evaluation, budgetary constraints, etc.).	Not Rated	Ibid
Question 23.	Does the evaluation meet UN SWAP evaluation performance indicators? (Note: this question will be rated according to UN SWAP standards)	67%	
	i GEEW is integrated in the Evaluation Scope of analysis, and evaluation criteria and questions are designed in a way that ensures GEEW-related data will be collected.	Satisfactorily integrated	GEEW is not highlighted in the scope or objectives of the evaluation, or in the main levels of enquiry in the learning framework used to guide the evaluation. However, there is a question specific to gender and equity.
	ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Fully integrated	The methodology was highly participatory. Data analysis processes are noted as disaggregating by gender. Efforts were made to ensure gender balance of respondents. Internal and external consultants, KII respondents, and workshop participants are disaggregated by gender and function - a reasonable level of gender balance was achieved.
	iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Partially integrated	Much of the enquiry was gender neutral/blind. GEEW is not reflected in the ToC developed as part of the DE. Gender analysis is apparent in one Finding (# 27 - innovation tools are not always designed with a lens of GEHR) but the discussion is very brief. There is no analysis of how men/women or girls/boys may experience or take up innovation differently. Gender is not reflected in Conclusions or 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 | <p>The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.</p> <ul style="list-style-type: none">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? |
|-----|---|