

GEROS Evaluation Quality Assurance Tool

Version 2016.4

Reviewers: complete all cells highlighted in Yellow

		Notes for QA
Title of the Evaluation Report	Evaluation of the Multilingual Education Component of the Unity in Diversity Project in Kyrgyzstan	
Report sequence number	2016/004	
Region	CEE/CIS	
Year of Report	2016	
Office	Rep of Kyrgyzstan	
Coverage (countries)	Rep of Kyrgyzstan	
ToRs present	Yes	ANNEX J, p97
Date of Review	1/11/2017	
Name of reviewer	ImpactReady (incl. Merits)	
Classification of Evaluation Report		
Management of Evaluation (<i>Managerial control and oversight of evaluation decisions</i>)		UNICEF managed
SPOA Correspondence <i>(Alignment with SPOA focus area priorities)</i>	Health	No
	HIV/AIDS	No
	WASH	No
	Nutrition	No
	Education	Yes
	Child protection	No
	Social inclusion	Yes
	Gender equality (cross-cutting)	Yes
	Humanitarian action (cross-cutting)	No
Evaluation object		Country Programme
Evaluation type		Formative
Evaluation strategy		Mixed methods
Evaluation design		Participatory

Evaluation level		Outcome
Geographic Scope		National
SECTION A: BACKGROUND (weight 5%)		Rating
Question 1.	Is the object of the evaluation clearly described?	89%
	Clear and relevant description of the intervention, including: location(s), timelines, cost/budget, and implementation status	Yes
	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/cities, sub-regions...) and in terms of numbers reached (as appropriate to the purpose of the evaluation)	Mostly
	Description of the relative importance of the object to UNICEF (e.g. in terms of size, influence, or positioning)	Yes
Question 2.	Is the context of the intervention clearly described?	100%
	Clear and relevant description of the context of the intervention (policy, socio-economic, political, institutional, international factors relevant to the implementation of the intervention)	Yes
	Clear and relevant description (where appropriate) of the status and needs of the target groups for the intervention	Yes
	Explanation of how the context relates to the implementation of the intervention	Yes
Question 3.	Is the results chain or logic well articulated?	83%
	Clear and complete description of the intervention's intended results	Yes
	Intervention logic presented as a coherent theory of change, logic chain or logic framework	Mostly
Question 4.	Are key stakeholders and their contributions clearly identified?	83%
	Identification of implementing agency(ies), development partners, primary duty bearers, secondary duty bearers, and rights holders	Mostly
	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF	Yes
•••-	Executive Feedback on Section A	Overall rating for Section
The rater will briefly (3-5 sentences) assess top line issues for this section relevant for feedback to senior management (positive and negative), summarizing here how the evaluation report meets or fails to meet the criteria above (judgement). As relevant, the rater will highlight best		Satisfactory
<p>The Evaluation object has been described clearly with objectives and relevance of the program. The context of the intervention has been well presented. Risk factors of the programme well articulated. Programme logic and a brief ToC has been given as specified in the ToRs. This could have been developed further. Programme budget and contributions of UNICEF and OHCHR_ROCA are given in Table 8 under Findings, but the overall amounts could have been mentioned in description of the intervention.</p>		 3
Additional comments for Section A (recommendations for improvement)		Weighting
Generally satisfactorily presented.		0.05
SECTION B: EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)		Rating
Question 5.	Is the purpose of the evaluation clearly described?	83%
	Specific identification of how the evaluation is intended to be used and to what this use is expected to achieve	Yes
	Identification of appropriate primary intended users of the evaluation	Mostly
Question 6.	Are the objectives and scope of the evaluation clear and realistic?	100%

	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	Yes
	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 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
••••	Executive Feedback on Section B	Overall rating for Section
	The rater will briefly (3-5 sentences) assess top line issues for this section relevant for feedback to senior management (positive and negative), summarizing here how the evaluation report meets or fails to meet the criteria above (judgement). As relevant, the rater will highlight best practice/added value elements	Satisfactory
	The section has been well written. The challenge in terms of language has been indicated. Primary and secondary users of the report could be identified for better utilisation of the report	 3
	Additional comments for Section B (recommendations for improvement)	Weighting
	The section generally meets the quality requirements	0.05

SECTION C: EVALUATION METHODOLOGY (weight 15%)		Rating
Question 7.	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 consider Coverage; Connectedness; Coordination; Protection; Security.	100%
	Clear and relevant presentation of the evaluation framework including clear evaluation questions used to guide the evaluation	Yes
	If the framework is OTHER than UNICEF standard criteria, or if not all standard criteria of the chosen framework are included, the reasons for this are clearly explained and the chosen framework is clearly described	Not Rated
Question 8.	Does the report specify methods for data collection, analysis, and sampling?	92%
	Clear and complete description of a relevant design and set of methods that are suitable for the evaluation's purpose, objectives and scope	Yes
	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
	Clear and complete description of the methods of analysis, including triangulation of multiple lines and levels of evidence (if relevant)?	Mostly
	Clear and complete description of limitations and constraints faced by the evaluation, including gaps in the evidence that was generated and mitigation of bias?	Yes
Question 9.	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:	44%
	Explicit reference to the obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability)	Partly
	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)	Mostly

	ONLY FOR THOSE CASES WHERE THE EVALUATION INVOLVES INTERVIEWING CHILDREN: explicit reference is made to the UNICEF procedures for Ethical Research Involving Children	Partly
•••-	Executive Feedback on Section C	Overall rating for Section
The rater will briefly (3-5 sentences) assess top line issues for this section relevant for feedback to senior management (positive and negative), summarizing here how the evaluation report meets or fails to meet the criteria above (judgement). As relevant, the rater will highlight best		Satisfactory
<p>Methodology for the evaluation has been explained. The evaluation methodology is generally as per ToRs and is stated to be according to UNICEF standard criteria. OECD-DAC evaluation criteria have been adopted. While it has been mentioned that data from various sources would be triangulated, specificity in the statement could have added value to the report. The lists of participants in the evaluation interviews and FGDs have been annexed. They could have been gender segregated. While there is a mention of ethical standards being followed, there is no specific indication how they are followed in practice</p>		 3
Additional comments for Section C (recommendations for improvement)		Weighting
The report would benefit from providing specific details on how the referenced ethics standards were applied.		0.15

SECTION D: EVALUATION FINDINGS (weight 20%)	Rating
--	---------------

Question 10.	Do the findings clearly address all evaluation objectives and scope?	83%
	Findings marshal sufficient levels of evidence to systematically address all of the evaluation's questions and criteria	Mostly
	Reference to the intervention's results framework in the formulation of the findings	Yes
Question 11.	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.	50%
	The evaluation clearly presents multiple lines (including multiple time series) and levels (output, outcome, and appropriate disaggregation) of credible evidence.	Mostly
	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.	Mostly
	Unexpected effects (positive and negative) are identified and analysed	No
	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).	Mostly
Question 12.	Does the evaluation assess and use the intervention's Results Based Management elements?	0%
	Clear and comprehensive assessment 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)	No
	Clear and complete assessment of the use of monitoring data in decision making	No

•••-	Executive Feedback on Section D	Overall rating for Section
The rater will briefly (3-5 sentences) assess top line issues for this section relevant for feedback to senior management (positive and negative), summarizing here how the evaluation report meets or fails to meet the criteria above (judgement). As relevant, the rater will highlight best		Fair

<p>The findings seem to be based on conclusions from FGDs and other stakeholder interviews. As a large number of stakeholders of different groups seem to have been covered in the evaluation, it would have made the evidence more credible if frequencies of different responses of the stakeholders were presented where appropriate and possible. The findings do not make a mention of the Monitoring system of the programme. Perhaps there is none.</p>	 <p>2</p>
--	--

<p>Additional comments for Section D (recommendations for improvement) UNICEF evaluations are expected to assess the extent to which a Results Based Management system was established and used to improve performance.</p>	<p>Weighting 0.2</p>
--	--------------------------

<p>SECTION E: EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 15%)</p>	<p>Rating</p>
--	---------------

<p>Question 13.</p>	<p>Do the conclusions present an objective overall assessment of the intervention?</p>	<p>89%</p>
---------------------	--	------------

	<p>Clear and complete description of the strengths and weaknesses of the intervention that adds insight and analysis beyond the findings</p>	<p>Yes</p>
--	--	-------------------

	<p>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)</p>	<p>Yes</p>
--	---	-------------------

	<p>The conclusions are derived appropriately from findings</p>	<p>Mostly</p>
--	--	----------------------

<p>Question 14.</p>	<p>Are lessons learned correctly identified?</p>	<p>67%</p>
---------------------	--	------------

	<p>Correctly 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.</p>	<p>Mostly</p>
--	---	----------------------

<p>• • •</p>	<p>Executive Feedback on Section E</p>	<p>Overall rating for Section</p>
--------------	---	-----------------------------------

<p>The rater will briefly (3-5 sentences) assess top line issues for this section relevant for feedback to senior management (positive and negative), summarizing here how the evaluation report meets or fails to meet the criteria above (judgement). As relevant, the rater will highlight best practice/added value elements.</p>	<p>Fair</p>
---	--------------------

<p>Lessons learnt and conclusions are clearly stated and are based on evaluation findings. However, the conclusions, placed in the report after the recommendations and lessons learnt are general . The section appears to be more generic</p>	 <p>2</p>
---	--

<p>Additional comments for Section E (recommendations for improvement) Conclusions should have been more specific abstractions from the findings and should have been placed before recommendations.</p>	<p>Weighting 0.15</p>
---	---------------------------

SECTION F: RECOMMENDATIONS (weight 15%)		Rating
Question 15.	Are recommendations well grounded in the evaluation?	78%
	Recommendations are logically derived from the findings and/or conclusions	Yes
	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
	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	Partly
Question 16.	Are recommendations clearly presented?	83%
	Clear identification of target group for action for each recommendation (or clearly clustered group of recommendations)	Yes
	Clear prioritization and/or classification of recommendations to support use	Mostly
•••-	Executive Feedback on Section F	Overall rating for Section
	The rater will briefly (3-5 sentences) assess top line issues for this section relevant for feedback to senior management (positive and negative), summarizing here how the evaluation report meets or fails to meet the criteria above (judgement). As relevant, the rater will highlight best practice/added value elements	Satisfactory
	The recommendations seem to be directly based on findings and have been neatly presented according to evaluation criteria, time-frame, and priority, indicating the organizations responsible.	 3
Additional comments for Section F (recommendations for improvement)		Weighting
This section makes specific recommendations and is well-presented. It would have been more logical to link the recommendations to the conclusions.		0.15

SECTION G: EVALUATION STRUCTURE/PRESENTATION (weight 5%)		Rating
Question 17.	Does the evaluation report include all relevant information?	100%
	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
	Annexes should include, when not present in the body of the report: Terms of Reference, Evaluation matrix, list of interviewees, list of site visits, data collection instruments (such as survey or interview questionnaires), list of documentary evidence Other appropriate annexes could include: additional details on methodology, copy of the results chain, information about the evaluator(s)	Yes
Question 18.	Is the report logically structured?	83%
	The structure is easy to identify and navigate (for instance, with numbered sections, clear titles and sub-titles)	Yes

	Context, purpose and methodology would normally precede findings, which would normally be followed by conclusions, lessons learned and recommendations	Mostly
•••-	Executive Feedback on Section G	Overall rating for Section
	The rater will briefly (3-5 sentences) assess top line issues for this section relevant for feedback to senior management (positive and negative), summarizing here how the evaluation report meets or fails to meet the criteria above (judgement). As relevant, the rater will highlight best	Satisfactory
	The structure of the report is generally logical. Conclusions could have come before Recommendations. Annexes could have included information on the evaluators.	 3
Additional comments for Section G (recommendations for improvement)		Weighting
Conclusions should come after findings but before recommendations.		0.05

SECTION H: EVALUATION PRINCIPLES (weight 15%)		Rating
Question 19.	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?	33%
	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	No
	Clear description of the level of participation of key stakeholders in the conduct of the evaluation, and description of the rationale for the chosen level of participation (for example, a reference group is established, stakeholders are involved as informants or in data gathering)	Mostly
	Stylistic evidence of the inclusion of these considerations 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.	Partly
Question 20.	Does the evaluation assess the extent to which the implementation of the intervention addressed gender, equity & child rights?	67%
	Identification and assessment of the presence or absence of equity considerations in the design and implementation of the intervention	Yes
	Identification and assessment of the presence or absence of gender in the design and implementation of the intervention	Partly
	Explicit analysis of the involvement in the object of right holders, duty bearers, and socially marginalized groups, and the differential benefits received by different groups of children	Mostly
	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.)	Mostly
Question 21.	Does the evaluation meet UN SWAP evaluation performance indicators?	Approaches requirements
	Note: this question will be rated according to UN SWAP standards	5
	GEEW is integrated in the Evaluation Scope of analysis and Indicators are designed in a way that ensures GEEW-related data will be collected	Partly integrated
	Evaluation Criteria and Evaluation Questions specifically address how GEEW has been integrated into the design, planning, implementation of the intervention and the results achieved.	Partly integrated

	A gender-responsive Evaluation Methodology, Methods and tools, and Data Analysis Techniques are selected.	Partially integrated 1 point
	The evaluation Findings, Conclusions and Recommendation reflect a gender analysis	Satisfactorily integrated 2 points
••••	Executive Feedback on Section H	Overall rating for Section
	The rater will briefly (3-5 sentences) assess top line issues for this section relevant for feedback to senior management (positive and negative), summarizing here how the evaluation report meets or fails to meet the criteria above (judgement). As relevant, the rater will highlight best practice/added value elements	Fair
	The ToRs call for covering equity and gender aspects in the evaluation. These aspects have been covered under Finding 4 in Section 4.1. However, there is not much evidence of gender and equity segregated data and analysis of results.	 2
Additional comments for Section H (recommendations for improvement)		Weighting
Gender and equity aspects should have been mainstreamed across the whole range of data collection tools, analyses, and recommendations.		0.15

SECTION I: EXECUTIVE SUMMARY (weight 5%)		Rating
Question 22.	Can the executive summary inform decision-making?	100%
	An executive summary is provided that is of relevant conciseness and depth for primary intended users	Yes
	Includes all necessary elements (overview of the intervention, evaluation purpose, objectives and intended audience, evaluation methodology, key findings, key conclusions, key recommendations)	Yes
	Includes all the necessary information to understand the intervention and the evaluation AND does not contain information not already included in the rest of the report	Yes
••••	Executive Feedback on Section I	Overall rating for Section
	The rater will briefly (3-5 sentences) assess top line issues for this section relevant for feedback to senior management (positive and negative), summarizing here how the evaluation report meets or fails to meet the criteria above (judgement). As relevant, the rater will highlight best	Satisfactory

<p>Executive summary has been well written indicating all the relevant points. It is well structured and logical.</p>	 <p style="text-align: right;">3</p>
---	--

<p>Additional comments for Section I (recommendations for improvement)</p> <p>Findings and recommendations have been neatly summarized. Only Conclusions could have come before recommendations.</p>	<p>Weighting</p> <p style="text-align: right;">0.05</p>
---	---

Weighting checksum (should equal 1) 1

OVERALL SCORE (max=4, min=0)

<p>• • • -</p>	<p>Satisfactory</p>	 <p style="text-align: right;">2.5</p>
<p>Meets UNICEF/UNEG standards for evaluation reports and decision makers may use the evaluation with confidence</p>		

Lessons for improving the management and performance of future evaluations

It is essential that the report should be seen to have utilized all the qualitative and quantitative data collected through detailed analysis of the evidence.