GEROS Evaluation Quality Assurance Tool

Version 2018.1

	Reviewers: complete all cells highlighted in Yellow	
		Notes for QA
Title of the Evaluation Report	**Evaluation of the Pilot MEND and RISE Cash Transfer Programmes in St. Kitts and Nevis**	
Report sequence number	**2018/001**	
Region	**LACR**	
Year of Report	**2018**	
Office	**Barbados**	
Coverage (countries)	**Barbados**	
ToRs present	**Yes**	
Date of Review	**7/13/2018**	
Name of reviewer	**ImpactReady**	

Classification of Evaluation Report		
Management of Evaluation *(Managerial control and oversight of evaluation decisions)*		**UNICEF managed**
SPOA Correspondence *(Alignment with SPOA focus area priorities)*	**Health**	**Yes**
	HIV/AIDS	**No**
	WASH	**No**
	Nutrition	**Yes**
	Education	**Yes**
	Child protection	**Yes**
	Social inclusion	**Yes**
	Gender equality (cross-cutting)	**Yes**
	Humanitarian action (cross-cutting)	**No**
Evaluation object		**Pilot/innovation**
Evaluation type		**Summative and formative**
Evaluation strategy		**Mixed methods**

Evaluation design		Quasi-experimental
Evaluation level		Impact
Geographic Scope		National

SECTION A: BACKGROUND (weight 5%)		100%
Question 1.	Is the object of the evaluation clearly described?	100%
	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/cites, sub-regions...) and in terms of numbers reached (as appropriate to the purpose of the evaluation)	Yes
	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?	100%
	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	Yes
Question 4.	Are key stakeholders and their contributions clearly identified?	100%
	Identification of implementing agency(ies), development partners, primary duty bearers, secondary duty bearers, and rights holders	Yes
	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 practice/added value elements	**Highly Satisfactory**
The report presents a thorough description of the sociopolitical context and the way it relates to the initiative. Similarly, the report clearly explains the object of the evaluation, including time and location, current status of the programme, primary beneficiaries as well as their main characteristics and needs. Furthermore, the report discusses the importance that the intervention has for UNICEF. A Theory of Change is discussed in detail and a graphic representation is presented . Finally, main stakeholders, their respective roles and contributions, along with their participation level in the evaluation process are duly discussed and presented in Table 4.	4

Additional comments for Section A (recommendations for improvement)	Weighting
This section observes good practices. No further improvement is required.	0.05

SECTION B: EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)		83%
Question 5.	Is the purpose of the evaluation clearly described?	100%
	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	**Yes**
Question 6.	Are the objectives and scope of the evaluation clear and realistic?	67%
	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)	**Partly**

• • • -	**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 purpose of the evaluation is clearly presented and appears realistic. Additionally, the report clearly indicates the evaluation's intended users. On the other hand, even if the report discusses the evaluation's main focus, it lacks detail around the scope of the evaluation in terms of time period and geographical coverage.		**3**
Additional comments for Section B (recommendations for improvement)		Weighting
It is important that the subsection dedicated to discussing the scope of the evaluation clearly addresses what is to be covered by the evaluation in chronological and geographic terms. Furthermore, the report should discuss what (if anything) was left out of the evaluation scope and provide a rationale for this.		0.05

SECTION C: EVALUATION METHODOLOGY (weight 15%)		83%
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	**Yes**
Question 8.	Does the report specify methods for data collection, analysis, and sampling?	100%

	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)?	**Yes**	
	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:		50%
	Explicit reference to the obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability)	**No**	
	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)	**Yes**	
	ONLY FOR THOSE CASES WHERE THE EVALUATION INVOLVES INTERVIEWING CHILDREN: explicit reference is made to the UNICEF procedures for Ethical Research Involving Children	**Not Rated**	

• • • -	**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 practice/added value elements		**Satisfactory**
The report presents a good discussion around the evaluation framework and the selection of evaluation criteria, which are aligned with the OECD/DAC standard criteria. The evaluation also includes cross-cutting principles of gender equality, equity and human rights according to the UNEG's Norms and Standards for Evaluation. The report also presents the methods of analysis used along with an explanation of how triangulation was conducted among different levels of evidence. Also, a detailed evaluation matrix with specific evaluation questions is presented as well as the limitations faced by the evaluation and some mitigation strategies. Finally, the evaluation does a good job at discussing ethical safeguards used to protect participants, including a discussion around the way UNICEF procedures for ethical research involving children were implemented. However, even if the report explains that evaluators employed a gender-sensitive approach that recognized the unique roles held by women and girls in creating social change, other ethical obligations of the evaluators are not discussed.		**3**

Additional comments for Section C (recommendations for improvement)	Weighting
It is recommended to explicitly address the obligations of the evaluators regarding their independence, impartiality, credibility, conflicts of interest, and accountability, as per UNEG's Norms and Standards. For more information on this, please visit http://www.unevaluation.org/document/detail/1914	0.15

SECTION D: EVALUATION FINDINGS (weight 20%)	100%
Question 10.	Do the findings clearly address all evaluation objectives and scope?	100%

	Findings marshal sufficient levels of evidence to systematically address all of the evaluation's questions and criteria	**Yes**

	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.		100%
	The evaluation clearly presents multiple lines (including multiple time series) and levels (output, outcome, and appropriate disaggregation) of credible evidence.		**Yes**
	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**
	Unexpected effects (positive and negative) are identified and analysed		**Yes**
	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**
Question 12.	Does the evaluation assess and use the intervention's Results Based Management elements?		100%
	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)		**Yes**
	Clear and complete assessment of the use of monitoring data in decision making		**Yes**

• • • •	**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 practice/added value elements		**Highly Satisfactory**
Findings follow the evaluation framework, respond to key evaluation questions, and make reference to outputs, short-term and long-term outcomes and impact. The evaluation addresses both positive and negative findings and they are supported by several lines of evidence providing causal factors that explain success or failure. Similarly, as per the ToRs, the evaluation report discusses the initiative's M&E system and the way this system was used during the implementation period, along with ways to improve it. Also, the report duly discusses the occurrence of unintended results and it explains that even if there were none, some results had higher importance than anticipated.		4
Additional comments for Section D (recommendations for improvement)		Weighting
This section observes good practices. No further improvement is required.		0.2

SECTION E: EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 15%)		100%

Question 13.	Do the conclusions present an objective overall assessment of the intervention?		100%
	Clear and complete description of the strengths and weaknesses of the intervention that adds insight and analysis beyond the findings		**Yes**
	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 derived appropriately from findings		**Yes**
Question 14.	Are lessons learned correctly identified?		100%
	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.		**Yes**

	Executive Feedback on Section E	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		**Highly Satisfactory**
Conclusions provide further insight into the information presented in the findings and they provide a discussion around the challenges that may lie ahead if the pilot programme is scaled-up. Futhermore, lessons learned are correctly identified and duly generalized so as to maximize their usefulness in other contexts.		4
Additional comments for Section E (recommendations for improvement)		Weighting
This section observes good practices. No further improvement is required.		0.15

SECTION F: RECOMMENDATIONS (weight 15%)		100%
Question 15.	Are recommendations well grounded in the evaluation?	100%
	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	**Yes**
Question 16.	Are recommendations clearly presented?	100%
	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	**Yes**
	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		**Highly Satisfactory**
Recommendations stem logically from the findings and conclusions and clearly indicate the target stakeholders for action in each case. Recommendations are clearly classified in a table format and cover programme areas, provide orders of priority and a rationale. Finally, the report duly describes the process followed in developing the recommendations and the participation of key stakeholders in this process.		4
Additional comments for Section F (recommendations for improvement)		Weighting

This section observes good practices. No further improvement is required.		0.15

SECTION G: EVALUATION STRUCTURE/PRESENTATION (weight 5%)	92%

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	**Mostly**
	Context, purpose and methodology would normally precede findings, which would normally be followed by conclusions, lessons learned and recommendations	**Yes**

• • • -	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 practice/added value elements		**Satisfactory**
The opening pages of the report include all necessary information in order to quickly grasp the essential elements of the intervention and the evaluation. Similarly, the annexes include valuable information such as methodological tools, the ToRs, the evaluators' biography, list of interviewees, etc. Finally, the evaluation report is easy to navigate, with clearly numbered sections and sub-sections following a logical structure. This being said, the evaluation report is too long (at 120 pages).		**3**

Additional comments for Section G (recommendations for improvement)	Weighting
This section generally observes good practices. However, the evaluation report is excessively long (at 120 pages) and could be shortened, for example, by annexing some information such as the evaluation matrix or the Theory of Change, etc. It is essential that the report be an accessible length to be useful to stakeholders (around 60 - 80 pages).	0.05

SECTION H: EVALUATION PRINCIPLES (weight 15%)	86%

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?	100%
	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**

	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)	**Yes**
	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.	**Yes**
Question 20.	Does the evaluation assess the extent to which the implementation of the intervention addressed gender, equity & child rights?	92%
	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	**Yes**
	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	**Yes**
	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	6
UN-SWAP criterion 1:	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 2 points**
UN-SWAP criterion 2:	A gender responsive **methodology, methods and tools, and data analysis** techniques are selected.	**Satisfactorily integrated 2 points**
UN-SWAP criterion 3:	The evaluation **findings, conclusions and recommendations** reflect a gender analysis.	**Satisfactorily integrated 2 points**
● ● ● -	**Executive Feedback on Section H**	Overall rating for Section
	The rater will briefly 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). The rater should include one labelled sentence for each of the UN-SWAP criteria.	**Satisfactory**
	The report uses human rights, equity, and gender sensitive language. It makes explicit reference to human rights obligations and standards, and breaks down rights holders per equity group. The report is strong at identifying stakeholders and at discussing their roles in the initiative as well as in the evaluation. However, stakeholder participation in the evaluation may have been even greater had the evaluation manager formed an Evaluation Reference Group made up of key stakeholders to help guide and manage the evaluation process. The evaluation made a strong attempt to address and mainstream gender equality. GEEW is mainstreamed through most (but not all) of the evaluation criteria and questions and is mentioned throughout the evaluation findings and recommendations. However, there remains room for the evaluation to have executed an even stronger gender focus.	**3**
	Additional justifications for UN-SWAP ratings	
UN-SWAP criterion 1:	**Gender Equality considerations are mainstreamed through some of the evaluation criteria but could be further mainstreamed (e.g. the experiences of boys and girls could have been broken down under questions of effectiveness).**	
UN-SWAP criterion 2:	**The report states that "throughout the research process, evaluators employed a gender-sensitive approach that recognized the unique roles held by women and girls in creating social change, as per United Nations guidelines". It would have been useful to the reader to provide a more detailed explanation as to how this was done.**	

UN-SWAP criterion 3:	Gender equality considerations are mentioned throughout the findings, conclusions, and recommendations. However, they could have been further mainstreamed had the evaluation matrix better insisted on disaggregating information by sex (e.g. instead of collecting information on "beneficiaries", it would have been useful to collect information on "boys and girls").	
Additional comments for Section H (recommendations for improvement)		Weighting
Stakeholder participation in the evaluation may have been even greater had the evaluation manager formed an Evaluation Reference Group made up of key stakeholders to help guide and manage the evaluation process. To learn more about conducting participatory evaluations, please see: http://www.betterevaluation.org/en/plan/approach/participatory_evaluation. There is also room for the evaluation to have used a stronger GEEW focus. For instance, GEEW could have been better mainstreamed through all of the evaluation questions and criteria by identifying stakeholders by sex. Another way of strengthening GEEW considerations could have been to use a specific gender equality evaluation crtierion in the evaluation matrix. This would likely have helped strengthen the report's mainstreaming of GEEW throughout the findings and recommendations. Finally, it would have been useful to the reader to understand exactly how a gender-sensitive approach was used within the evaluation methodology. To learn more about how to conduct gender sensitive evaluations, please see: http://www.betterevaluation.org/en/resources/guide/gender-responsive_evaluation_handbook.		0.15

SECTION I: EXECUTIVE SUMMARY (weight 5%)		89%
Question 22.	Can the executive summary inform decision-making?	89%
	An executive summary is provided that is of relevant conciseness and depth for primary intended users	**Mostly**
	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 practice/added value elements		**Satisfactory**
The executive summary provides all of the necessary elements to allow for a thorough understanding of both the initiative as well as the evaluation. However, the executive summay is excessively lengthy (12 pages) and should be further synthetized to 4-5 pages to be useful to decision-makers. Finally, all elements addressed in the executive summary are discussed in further detail in the body of the report.		3
Additional comments for Section I (recommendations for improvement)		Weighting
The executive summary for the most part observes good practices. It is recommended that it be shortened to 4-5 pages in order to successfully and succinctly inform end-users.		0.05

Weighting checksum (should equal 1) 1

OVERALL SCORE (max=4, min=0)		
• • • • **Highly Satisfactory**		**3.55**

Exceeds UNICEF/UNEG standards for evaluation reports and decision makers may use the evaluation with a high degree of confidence
Lessons for improving the management and performance of future evaluations
This is a strong evaluation report that is based on a good methodology and that presents very robust and well written findings, conclusions, and recommendations. The report made a good attempt at integrating and mainstreaming gender equality and at using a paritcipatory approach. However, the evaluation could have been even more participatory had the evaluation manager set up an Evaluation Reference Group made up of key stakeholders to help guide and manage the evaluation. Also, there is room to more thoroughly address GEEW throughout the evaluation process and report. In the future, it would be good for the evaluation manager to specify that the report should be no longer than 80 pages to ensure that it is accessible and can be easily used by stakeholders.