EVALUATION ID					
	4550-2015/008				

UNICEF Global Evaluation Report Oversight System (GEROS) Review Template

	CC	Dark green	Green	Amber	Red	White
Colour Coding	Questions	Outstanding	Yes	Mostly Satisfactory	No	Not Applicable
	Section & Overall Rating	Outstanding, best practice	Highly Satisfactory	Mostly Satisfactory	Unsatisfactory	

The Cornerstone questions are in column J and are questions that need to be answered for rating and justification of each of the six sections

[UNEG Standards for Evaluation in the UN System](#)	[UNEG Norms for Evaluation in the UN System](#)	[UNICEF Adapted UNEG Evaluation Report Standards](#)

	Response
Title of the Evaluation Report	**Evaluation of TUSEME Programme in Tanzania**

Report sequence number	2015/008	**Date of Review**	2015/12/31	**Year of the Evaluation Report**	2015
Region	Eastern and Southern Africa Regional Office			**Country**	United Rep. of Tanzania
Type of Report	Evaluation			**TORs Present**	Yes
Name of reviewer	Universalia Management Group				

Classification of Evaluation Report		Comments
Geographic Scope (*Coverage of the programme being evaluated & generalizability of evaluation findings*)	1.1 Sub-national: The programme and evaluation covers selected sub-national units (districts, provinces, states, etc.) within a country, where results cannot be generalized to the whole country	
Management of Evaluation (*Managerial control and oversight of evaluation decisions*)	2.1 UNICEF managed: Working with national partners of different categories UNICEF is responsible for all aspects of the evaluation.	
Purpose (*Speaks to the overarching goal for conducting the evaluation; its raison d'être*)	3.7 Programme: An evaluation of a sectorial programme to determine its overall effectiveness and efficiency in relation to the stated goals and objectives	
Result (*Level of changes sought, as defined in RBM: refer to substantial use of highest level reached*)	4.3 Impact: Final results of a programme or policy on the intended beneficiaries and, where possible, on comparison groups. Reflects the cumulative effect of donor supported programmes of cooperation and national policy initiatives.	
SPOA Correspondence (*Alignment with SPOA focus area priorities: (1) Health; (2) HIV-AIDS; (3) WASH; (4) Nutrition; (5) Education; (6) Child Protection; (7) Social Inclusion; (8) Cross-Cutting - Gender Equality; and (9) Cross-cutting - Humanitarian Action)*)	5.8 Cross cutting - Gender Equality: UNICEF will emphasize the empowerment of girls and women and address gender-related needs of girls, boys, fathers, mothers and communities. UNICEF will identify and leverage positive cross-sectoral synergies and linkages such as those among improving girls' education, ending child marriage and reducing maternal mortality. UNICEF will also focus on increasing access to services and opportunities by women and girls and their inclusion and participation in all facets of life as well as on advocacy and technical support on gender-equitable policies, budgeting and resource allocations. Emphasis will be placed on: a) Sex-disaggregated and other gender related data. b) Promote gender-sensitive interventions as a core programmatic priority, c) To the extent possible, all relevant policies, programmes and activities will mainstream gender equality.	
Level of Independence (*Implementation and control of the evaluation activities*)	6.3 Independent external: The evaluation is implemented by external consultants and/or UNICEF Evaluation Office professionals. The overall responsibility for the evaluation lies outside the division whose work is being evaluated.	
Approach	7.3 Summative and formative: An evaluation that combines the elements of a formative and a summative evaluation.	

SECTION A: OBJECT OF THE EVALUATION

Question	cc	Remarks	A/ Does the report present a clear & full description of the 'object' of the evaluation?	Constructive feedback for future reports
Object and context				
1 Is the object of the evaluation well described? This needs to include a clear description of the interventions (project, programme, policies, otherwise) to be evaluated including how the designer thought that it would address the problem identified, implementing modalities, other parameters including costs, relative importance in the organization and (number of) people reached.	Mostly	The Object of evaluation is the TUSEME Programme implemented in the United Rep. of Tanzania and it is described on pages 11-14. The interventions to be evaluated were the TUSEME Clubs, the degree of student participation in them and the impact they have had on girls' empowerment and the acquisition of life skills (pages 12-13). On page 12, the report identifies the number of teachers and students reached by the programme. However, the scope of the programme in terms of time period and the programming period being evaluated is not clear. Additionally, there is no financial information related to the programme provided. There is a section dedicated to describing the context, which includes information on the social, demographic and economic situation in Tanzania (pages 9-11). However, this section does not include information on women's rights or national policies that address gender equality, which are issues that lay at the heart of the TUSEME programme. Therefore, the contextual information is not as useful as it could be in terms of illuminating the findings.	**A/ Does the report present a clear & full description of the 'object' of the evaluation?** The report should describe the object of the evaluation including the results chain, meaning the 'theory of change' that underlies the programme being evaluated. This theory of change includes what the programme was meant to achieve and the pathway (chain of results) through which it was expected to achieve this. The context of key social, political, economic, demographic, and institutional factors that have a direct bearing on the object should be described. For example, the partner government's strategies and priorities, international, regional or country development goals, strategies and frameworks, the concerned agency's corporate goals & priorities, as appropriate.	**Constructive feedback for future reports** *Including how to address weaknesses and maintaining good practice*
2 Is the context explained and related to the object that is to be evaluated? The context includes factors that have a direct bearing on the object of the evaluation: social, political, economic, demographic, institutional. These factors may include strategies, policies, goals, frameworks & priorities at the: international level; national Government level; individual agency level	Mostly			
3 Does this illuminate findings? The context should ideally be linked to the findings so that it is clear how the wider situation may have influenced the outcomes observed.	Mostly			
Theory of Change			The object of the evaluation and the nature of the programme's interventions as well as the context in which they took place are described but there is overall a lack of key information regarding these important elements of the evaluation. Namely, the report does not include contextual information on women's rights or national policies that address gender equality. Although the evolution of the programme is well explained, it is unclear what time period the evaluation is looking at. Also, there is no financial information on the programme provided. A Theory of Change model based on the Gender Responsive Model is discussed. However, it does not include a results chain. The key stakeholders and their respective contributions, financial and others, are not very clearly identified or described. Although the evolution of the programme is well described, there is no mention of changes to plans or strategies.	Some key information is missing in this evaluation which compromises the clarity of the report. The scope in terms of the time period that the evaluation is assessing should be clearly stated. Although the report provides relevant information on the social and economic context in Tanzania, contextual information on issues relating to the programme (i.e. women's rights and gender equality) should be discussed in order to help illuminate the findings. Although the description of the Theory of Change and the reasons for having chosen it are well described and justified, there should also be a results chain provided that links outputs to outcomes and impact. The key stakeholders and their contributions need to be clearly identified and described. Although the evolution of the programme is well described, the evaluation should also mention changes to
4 Is the results chain or logic well articulated? The report should identify how the designers of the evaluated object thought that it would address the problem that they had identified. This can include a results chain or other logic models such as theory of change. It can include inputs, outputs and outcomes, it may also include impacts. The models need to be clearly described and explained.	Mostly	A Theory of Change is discussed on pages 13-14 and it is based on a Gender Responsive Model (GRM), which advocates for a rights-based approach. Its use is justified as it is a participatory approach in which children are viewed as active observers and teachers are requested to be sensitive to gender-related barriers to learning among girls and boys. The Theory of Change does not, however, include a results chain linking outputs to outcomes and impact.		
Stakeholders and their contributions				
5 Are key stakeholders clearly identified? These include o implementing agency(ies) o development partners o rights holders o primary duty bearers o secondary duty bearers	Mostly	Key stakeholders are mentioned sporadically throughout the report but they are not clearly identified and their contributions are not well described. Some information on stakeholder contributions to the programme throughout the programme's development can be found on pages 11-12. Likewise, the		
6 Are key stakeholders' contributions described? This can involve financial or other contributions and should be specific. If joint program also specify UNICEF contribution, but if basket funding question is not applicable	Mostly			

					plans or strategies.
7 Are UNICEF contributions described? This can involve financial or other contributions and should be specific	Mostly	contributions of UNICEF are generally described on page 12 but lack specificity. Specific information on human resources or financial contributions is missing.	Mostly Sat		

Implementation Status

8 Is the implementation status described? This includes the phase of implementation and significant changes that have happened to plans, strategies, performance frameworks, etc that have occurred - including the implications of these changes	Mostly	The implementation status of the programme along with a description of its evolution is provided on page 12. Since 2004, more than 600 secondary school teachers and about 17,600 secondary school students have been trained on the TUSEME approach. Furthermore, Appendix D, on page 64 provides a good summary that describes the implementation process. The evolution of the programme is well described on pages 11-12. However, no mention is made of changes to plans, strategies, etc that may have occurred during the programme implementation.		

Executive Feedback on Section A Issues for this section relevant for feedback to senior management (positives & negatives), & justify rating. *Up to two sentences*	**Information around the programme context does not directly relate to the evaluated object, thus reducing its usefulness. Although the programme's Theory of Change is somewhat discussed, no results chain is provided.**

SECTION B: EVALUATION PURPOSE, OBJECTIVES AND SCOPE

Question	cc	Remarks		
Purpose, objectives and scope				
9 Is the purpose of the evaluation clear? This includes why the evaluation is needed at this time, who needs the information, what information is needed, how the information will be used.	Mostly	The purpose of the evaluation is described on page 14 and it entails examining the effectiveness of the implementation of the TUSEME programme and the emerging impact of the project. It is said that the evaluation is to provide overall advice to UNICEF about the pros and cons of the programme. This description along with a list of evaluation objectives and evaluation questions are presented on page 15. However, the section does not provide information about how the evaluation will be used and what specific purpose it will serve, i.e. scaling up the programme or providing accountability to funders, etc. Although geographic scope is mentioned, the scope of the evaluation concerning the time frame is not clearly specified.	**B/ Are the evaluation's purpose, objectives and scope sufficiently clear to guide the evaluation?** The purpose of the evaluation should be clearly defined, including why the evaluation was needed at that point in time, who needed the information, what information is needed, and how the information will be used. The report should provide a clear explanation of the evaluation objectives and scope including main evaluation questions and describes and justifies what the evaluation did and did not cover. The report should describe and provide an explanation of the chosen evaluation criteria, performance standards, or other criteria used by the evaluators.	**Constructive feedback for future reports** *Including how to address weaknesses and maintaining good practice*
10 Are the objectives and scope of the evaluation clear and realistic? This includes: Objectives should be clear and explain what the evaluation is seeking to achieve; Scope should clearly describe and justify what the evaluation will and will not cover; Evaluation questions may optionally be included to add additional details	Mostly			
11 Do the objective and scope relate to the purpose? The reasons for holding the evaluation at this time in the project cycle (purpose) should link logically with the specific objectives the evaluation seeks to achieve and the boundaries chosen for the evaluation (scope)	Yes			
Evaluation framework				
12 Does the evaluation provide a relevant list of evaluation criteria that are explicitly justified as appropriate for the Purpose? It is imperative to make the basis of the value judgements used in the evaluation transparent if it is to be understood and convincing. UNEG evaluation standards refer to the OECD/DAC criteria, but other criteria can be used such as Human rights and humanitarian criteria and standards (e.g. SPHERE Standards) but this needs justification.. Not all OECD/DAC criteria are relevant to all evaluation objectives and scopes. The TOR may set the criteria to be used, but these should be (re)confirmed by the evaluator. Standard OECD DAC Criteria include: Relevance; Effectiveness; Efficiency; Sustainability; Impact Additional humanitarian criteria include; Coverage; Coordination; Coherence; Protection; timeliness; connectedness; appropriateness. *(This is an extremely important question to UNICEF)*	Yes	The evaluation criteria are clearly presented on page 14 and further explained along with some key evaluation questions in Table 1, on page 15. They are consistent with the OECD/DAC criteria, namely; Relevance, Effectiveness, Efficiency, Impact and Sustainability.	**Mostly Satisfactory** The purpose of the evaluation does not provide information about the manner in which the evaluation will be used and what specific purpose it will serve. Also, the scope of the evaluation concerning the time frame is not clearly specified. The evaluation criteria are clearly described and discussed in the report and they are consistent with the OECD/DAC criteria.	The report should explain the way in which the evaluation will be used and its specific purpose. The scope of the evaluation concerning the time period should be clearly stated.
13 Does the evaluation explain why the evaluation criteria were chosen and/or any standard DAC evaluation criteria (above) rejected? The rationale for using each particular non-OECD-DAC criterion (if applicable) and/or rejecting any standard OECD-DAC criteria (where they would be applicable) should be explained in the report.	Yes			
Executive Feedback on Section B Issues for this section relevant for feedback to senior management (positives & negatives), & justify rating. *Up to two sentences*		**The evaluation is strong at describing, discussing, and explicitly justifying the selected evaluation criteria. However, the specific purpose of the evaluation and how it will be used is unclear as is the time period that the evaluation is covering.**		

SECTION C: EVALUATION METHODOLOGY, GENDER, HUMAN RIGHTS AND EQUITY

Question	cc	Remarks		
Data collection				
14 Does the report specify data collection methods, analysis methods, sampling methods and benchmarks? This should include the rationale for selecting methods and their limitations based on commonly accepted best practice.	Yes	The Evaluation Methodology is discussed in detail from pages 16-21. The need to collect quantitative and qualitative data is stated for triangulation purposes. The 12 schools and respondents chosen for the evaluation and the rationale are described on page 16. Pupils and education officials were chosen as participants for a diversity of perspectives. A detailed list of the respondents used in the evaluation is presented on Table 2 (page 17). Table 3 on page 18 provides a good account of the type of data collected, the source and the tool used in each case. The steps followed for Data Analysis are thoroughly described on pages 19-20. Methodological Limitations are discussed on page 22.	**C/ Is the methodology appropriate and sound?** The report should present a transparent description of the methodology applied to the evaluation that clearly explains how the evaluation was specifically designed to address the evaluation criteria, yield answers to the evaluation questions and achieve the evaluation purposes. The report should also present a sufficiently detailed description of methodology in which methodological choices are made explicit and justified and in which limitations of methodology applied are included. The report should give the elements to assess the appropriateness of the methodology. Methods as such are not 'good' or 'bad', they are only so in relation to what one tries to get to know as part of an evaluation. Thus this standard assesses the suitability of the methods selected for the specifics of the evaluation concerned, assessing if the methodology is suitable to the subject matter and the information collected are sufficient to meet the evaluation objectives.	**Constructive feedback for future reports** *Including how to address weaknesses and maintaining good practice*
15 Does the report specify data sources, the rationale for their selection, and their limitations? This should include a discussion of how the mix of data sources was used to obtain a diversity of perspectives, ensure accuracy & overcome data limits	Yes			
Ethics				
16 Are ethical issues and considerations described? The design of the evaluation should contemplate: How ethical the initial design of the programme was; The balance of costs and benefits to participants (including possible negative impact) in the programme and in the evaluation; The ethics of who is included and excluded in the evaluation and how this is done	No	The objective of the programme is ethical in its nature due to its mandate to promote gender equality. However, the evaluation does not specifically address any ethical considerations related to the evaluation's design and implementation. A sub-section is dedicated to Ethical Considerations, such as the confidentiality of the information provided by respondents and the voluntary basis of their participation (page 21). Also, the sub-section discusses consent for minors participating in the evaluation process.		
17 Does the report refer to ethical safeguards appropriate for the issues described? When the topic of an evaluation is contentious, there is a heightened need to protect those participating. These should be guided by the UNICEF Evaluation Office Technical Note and include: protection of confidentiality; protection of rights; protection of dignity and welfare of people (especially children); Informed consent; Feedback to participants; Mechanisms for shaping the behaviour of evaluators and data collectors	Yes			
Results Based Management				
18 Is the capability and robustness of the evaluated object's monitoring system adequately assessed? The evaluation should consider the details and overall functioning of the management system in relation to results: from the M&E system design, through individual tools, to the use of data in management decision making.	Yes	The weaknesses of the programme's M&E system are abundantly mentioned throughout the report and are included as one of the evaluation's Methodological Limitations on page 22. It is explicitly stated that the analysis and interpretation of almost	The Evaluation methodology as well as the data collection and analysis methods are well described. Data sources and collection tools are also clearly identified. The evaluation limitations include the absence of baseline data, benchmarks, and M&E information; the decision to not include a	Further actions by the evaluators should have been carried out to overcome limitations in order for the evaluation findings to be more authoritative. Issues relating to equity should also be mainstreamed.

19 Does the evaluation make appropriate use of the M&E framework of the evaluated object?

In addition to articulating the logic model (results chain) used by the programme, the evaluation should make use of the object's logframe or other results framework to guide the assessment. The results framework indicates how the programme design team expected to assess effectiveness, and it forms the guiding structure for the management of implementation.

Mostly

all evaluation criteria was affected by these limitations. The results framework of the evaluation is structured so as to assess the effectiveness of the programme in providing a platform for young people to speak and express views about the problems they face and how to address them. This is well described on pages 26-32. However, findings are not clearly structured around the programme's 5 objectives.

strong counterfactual methodology; missing financial information; and the inability to observe TUSEME clubs in action. These are significant limitations and the report suggests that they may have compromised the quality of the evaluation findings. Ethical considerations related to the behaviour of the evaluation team is discussed. However, ethical considerations around the programme's design and implementation are not assessed. The weaknesses of the programme's M&E system are abundantly mentioned

Human Rights, Gender and Equity				
20 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? This could be done in a variety of ways including: use of a rights-based framework, use of CRC, CCC, CEDAW and other rights related benchmarks, analysis of right holders and duty bearers and focus on aspects of equity, social exclusion and gender. Style includes: using human-rights language; gender-sensitive and child-sensitive writing; disaggregating data by gender, age and disability groups; disaggregating data by socially excluded groups. Promote gender-sensitive interventions as a core programmatic priority, To the extent possible, all relevant policies, programmes and activities will mainstream gender equality.	Yes	The evaluation uses the Gender Responsive Model which is a rights-based approach. Additionally, human-rights language is used throughout the evaluation and specifically to explain the Theory of Change (pages 13-14). Findings, Conclusions, Recommendations and Lessons Learned all include sections or comments on girls' rights and empowerment (pages 14, 26, 40, 43). The evaluation does a good job at discussing how the programme's activities had an impact on boys as well as on girls. Since the M&E system was lacking in most of the activities implemented by the programme, there is little information concerning gender monitoring processes. There is no explicit mention of or consideration made regarding equity in the evaluation.		throughout the report. The evaluation is strong at mainstreaming gender and human rights but does not discuss issues around equity.
21 Does the evaluation assess the extent to which the implementation of the evaluated object was monitored through human rights (inc. gender, equity & child rights) frameworks? UNICEF commits to go beyond monitoring the achievement of desirable outcomes, and to ensure that these are achieved through morally acceptable processes. The evaluation should consider whether the programme was managed and adjusted according to human rights and gender monitoring of processes.	Mostly			
22 Do the methodology, analytical framework, findings, conclusions, recommendations & lessons provide appropriate information on HUMAN RIGHTS (inc. women & child rights)? The inclusion of human rights frameworks in the evaluation methodology should continue to cascade down the evaluation report and be obvious in the data analysis, findings, conclusions, any recommendations and any lessons learned. If identified in the scope the methodology should be capable of assessing the level of: Identification of the human rights claims of rights-holders and the corresponding human rights obligations of duty-bearers, as well as the immediate underlying & structural causes of the non realisation of rights.; Capacity development of rights-holders to claim rights, and duty-bearers to fulfil obligations. Support for humanitarian action – achieving faster scaling up of response, early identification of priorities and strategies, rapid deployment of qualified staff and clear accountabilities and responses consistent with humanitarian principles in situations of unrest or armed conflict.	Yes		Mostly Satisfactory	
23 Do the methodology, analytical framework, findings, conclusions, recommendations & lessons provide appropriate information on GENDER EQUALITY AND WOMEN'S EMPOWERMENT? The inclusion of gender equality frameworks in the evaluation methodology should continue to cascade down the evaluation report and be obvious in the data analysis, findings, conclusions, any recommendations and any lessons learned. If identified in the scope the methodology should be capable of assessing the immediate underlying & structural causes of social exclusion; and capacity development of women to claim rights, and duty-bearers to fulfil their equality obligations.	Yes			
24 Do the methodology, analytical framework, findings, conclusions, recommendations & lessons provide appropriate information on EQUITY? The inclusion of equity considerations in the evaluation methodology should continue to cascade down the evaluation report and be obvious in the data analysis, findings, conclusions, any recommendations and any lessons learned. If identified in the scope the methodology should be capable of assessing the capacity development of rights-holders to claim rights, and duty-bearers to fulfil obligations & aspects of equity.	No			

Stakeholder participation			
25 Are the levels and activities of stakeholder consultation described? This goes beyond just using stakeholders as sources of information and includes the degree of participation in the evaluation itself. The report should include the rationale for selecting this level of participation. Roles for participation might include: o Liaison o Technical advisory o Observer o Active decision making The reviewer should look for the soundness of the description and rationale for the degree of participation rather than the level of participation itself.	Yes	The report is strong at describing the participation of stakeholders in the creation and testing of data collection tools (pages 19-21). On page 20, triangulation of data methods as well as pilot tested data collection tools prior to fieldwork are well described. It appears as though the selected level of stakeholder participation was appropriate.	
26 Are the levels of participation appropriate for the task in hand? The breadth & degree of stakeholder participation feasible in evaluation activities will depend partly on the kind of participation achieved in the evaluated object. The reviewer should note here whether a higher degree of participation may have been feasible & preferable.	Yes		
Methodological robustness			
27 Is there an attempt to construct a counterfactual or address issues of contribution/attribution? The counterfactual can be constructed in several ways which can be more or less rigorous. It can be done by contacting eligible beneficiaries that were not reached by the programme, or a theoretical counterfactual based on historical trends, or it can also be a comparison group.	N/A	Although the inclusion of pupils who were not reached by the TUSEME programme made part of the TORs, at a feedback meeting with UNICEF staff it was decided not to include schools without TUSEME clubs in data collection. This is discussed as part of the Methodological Limitations on page 22. Nevertheless, there are some comparisons made with students who did not participate in the TUSEME Clubs (pages 29-31) without much explanation as to how this data was collected. The Methodology used in the evaluation is the same as the model used by the TUSEME programme, i.e. the GRM, and it relates in a straightforward manner to the Purpose of the evaluation, as described on pages 14-15. Methodological Limitations are discussed on page 22 and issues of potential biases are discussed. Limitations include the absence of baseline data, benchmarks, and M&E information; the decision to not include a strong counterfactual methodology; missing financial information; and the inability to observe TUSEME clubs in action. The report states that "this evaluation has only established associations between TUSEME and some observed positive attributes in the target audience as reported by the respondents without establishing a cause-effect relationship" (p. 22). The evaluation limitations appear to be somewhat unacceptable due to the extent to which they have compromised the evaluation findings.	
28 Does the methodology facilitate answers to the evaluation questions in the context of the evaluation? The methodology should link back to the Purpose and be capable of providing answers to the evaluation questions.	Mostly		
29 Are methodological limitations acceptable for the task in hand? Limitations must be specifically recognised and appropriate efforts taken to control bias. This includes the use of triangulation, and the use of robust data collection tools (interview protocols, observation tools etc). Bias limitations can be addressed in three main areas: Bias inherent in the sources of data; Bias introduced through the methods of data collection; Bias that colours the interpretation of findings	Mostly		

Executive Feedback on Section C Issues for this section relevant for feedback to senior management (positives & negatives), & justify rating. *Up to two sentences*	**Methodological limitations are well discussed and include the absence of baseline data, benchmarks, and M&E information; the decision to not include a strong counterfactual methodology; missing financial information; and the inability to observe TUSEME clubs in action. These are serious limitations that appear to have compromised the validity of the evaluation findings.**

SECTION D: FINDINGS AND CONCLUSIONS

Question	cc	Remarks	D/ Are the findings and conclusions, clearly presented, relevant and based on evidence & sound analysis? Findings should respond directly to the evaluation criteria and questions detailed in the scope and objectives section of the report. They should be based on evidence derived from data collection and analysis methods described in the methodology section of the report. Conclusions should present reasonable judgments based on findings and substantiated by evidence, providing insights pertinent to the object and purpose of the evaluation.	Constructive feedback for future reports *Including how to address weaknesses and maintaining good practice*
Completeness and logic of findings				
30 Are findings clearly presented and based on the objective use of the reported evidence? Findings regarding the inputs for the completion of activities or process achievements should be distinguished clearly from results. Findings on results should clearly distinguish outputs, outcomes and impacts (where appropriate). Findings must demonstrate full marshalling and objective use of the evidence generated by the evaluation data collection. Findings should also tell the 'whole story' of the evidence and avoid bias.	Yes	Findings are clearly presented and thoroughly discussed through pages 24-37. Furthermore, the Findings address all of the evaluation criteria. Findings demonstrate a logical chain of progression from implementation to results. The latter is particularly well explained on pages 26-29 by discussing the platform provided by TUSEME to the youth so that they could express and address the problems they face and by doing so, especially empower girls who, according to evidence presented on pages 26-28, were in a better position to take action to counter abuse and boost their self-esteem, as shown by data in figure 1 on page 29. Gaps and limitations including missing baseline data, a poor M&E system, and missing information relating to financial flows are discussed and their implications recognised on page 22. Concerning unexpected findings, the overshadowing of other clubs existing at some schools by TUSEME Clubs due to the good functioning and reliability of the latter is discussed on pages 36-37 as well as the expansion of the issues the programme was initially expected to address. This unexpected finding is mentioned in the recommendations on page 45. Another unexpected finding is the drop of teachers' absenteeism to TUSEME Clubs, mentioned on page 32, as pupils become more confident to report poor behaviour and defend their rights.	Findings are clearly presented and thoroughly discussed. Findings adequately respond to all evaluation criteria and demonstrate a logical chain of progression from implementation to results. Gaps and limitations in the evaluation are discussed and the reader is guided on how to interpret the findings based on these limitations. Although a Cost Analysis is included in the TORs and also in the Evaluation Criteria, its assessment was in the end unattainable since the evaluator did not have the necessary financial information to do so. The evaluation assigns contribution to results and discusses causality. The conclusions are addressed to the corresponding stakeholders and provide useful insightful, and clearly identified key elements based on the findings. However, the conclusions do not present a discussion around underlying problems and key priority areas, as they only mention the programme's strengths.	Some kind of Cost Analysis should have been provided in the report as it was included in the TORs and mentioned in the Evaluation Criteria. Conclusions should address the programme's weaknesses and not only its strengths.
31 Do the findings address all of the evaluation's stated criteria and questions? The findings should seek to systematically address all of the evaluation questions according to the evaluation framework articulated in the report.	Yes			
32 Do findings demonstrate the progression to results based on the evidence reported? There should be a logical chain developed by the findings, which shows the progression (or lack of) from implementation to results.	Yes			
33 Are gaps and limitations discussed? The data may be inadequate to answer all the evaluation questions as satisfactorily as intended, in this case the limitations should be clearly presented and discussed. Caveats should be included to guide the reader on how to interpret the findings. Any gaps in the programme or unintended effects should also be addressed.	Yes			
34 Are unexpected findings discussed? If the data reveals (or suggests) unusual or unexpected issues, these should be highlighted and discussed in terms of their implications.	Yes			
Cost Analysis				
35 Is a cost analysis presented that is well grounded in the findings reported? Cost analysis is not always feasible or appropriate. If this is the case then the reasons should be explained. Otherwise the evaluation should use an appropriate scope and methodology of cost analysis to answer the following questions: o How programme costs compare to other similar programmes or standards o Most efficient way to get expected results o Cost implications of scaling up or down o Cost implications for replicating in a different context o Is the programme worth doing from a cost perspective o Costs and the sustainability of the programme.	No	Although a cost analysis is included in the TORs and in the Evaluation Criteria as part of the Efficiency criterion, the assessment of this is largely inconclusive as it was impossible for the evaluator to track the flow of resources, as explained on pages 33-34.		

Contribution and causality		
36 Does the evaluation make a fair and reasonable attempt to assign contribution for results to identified stakeholders? For results attributed to the programme, the result should be mapped as accurately as possible to the inputs of different stakeholders.	Yes	The success of the TUSEME programme is largely attributed to the passion and positive attitude of teachers involved in the programme, as well as the training of teachers and DEOs in a Gender Responsive Approach, thus motivating them to adapt an attitudinal change towards girls (page 40).
37 Are causal reasons for accomplishments and failures identified as much as possible? These should be concise and usable. They should be based on the evidence and be theoretically robust. *(This is an extremely important question to UNICEF)*	Yes	
Strengths, weaknesses and implications		
38 Are the future implications of continuing constraints discussed? The implications can be, for example, in terms of the cost of the programme, ability to deliver results, reputational risk, and breach of human rights obligations.	Yes	There is quite a complete section entitled "Challenges" facing the implementation of the TUSEME project on pages 37-39 that provides a number of useful insights about elements that need to be addressed so as to avoid failure or shortcomings in the future implementation of the programme. The Conclusions are presented on page 42-43 and they provide a summary of the strengths of the programme though not much is said about its weaknesses in this section. Among the strengths, the programme proves to be relevant, the participatory approach is its hallmark, and it presents unintended positive effects.
39 Do the conclusions present both the strengths and weaknesses of the evaluated object? Conclusions should give a balanced view of both the stronger aspects and weaker aspects of the evaluated object with reference to the evaluation criteria and human rights based approach.	Mostly	
Completeness and insight of conclusions		
40 Do the conclusions represent actual insights into important issues that add value to the findings? Conclusions should go beyond findings and identify important underlying problems and/or priority issues. Simple conclusions that are already well known do not add value and should be avoided.	Mostly	The Conclusions provide useful, insightful, and clearly identified key elements informed by the findings, such as the central role played by a participatory approach in the success of the programme; the added value of unintended achievements like the improvement in teachers' overall attendance; and the complementarity of TUSEME and the national development policy frameworks. Conclusions successfully take into consideration the view of all stakeholders (page 43). The evaluation demonstrates good practices in the way its Conclusions are addressed to all stakeholders and participants. Likewise, other users can also benefit from the Conclusions as they are presented in a clear language that is accessible to everyone. However, the Conclusions do not present a discussion around underlying problems and key priority areas because they only address the programme's strengths.
41 Do conclusions take due account of the views of a diverse cross-section of stakeholders? As well as being logically derived from findings, conclusions should seek to represent the range of views encountered in the evaluation, and not simply reflect the bias of the individual evaluator. Carrying these diverse views through to the presentation of conclusions (considered here) is only possible if the methodology has gathered and analysed information from a broad range of stakeholders.	Yes	
42 Are the conclusions pitched at a level that is relevant to the end users of the evaluation? Conclusions should speak to the evaluation participants, stakeholders and users. These may cover a wide range of groups and conclusions should thus be stated clearly and accessibly: adding value and understanding to the report (for example, some stakeholders may not understand the methodology or findings, but the conclusions should clarify what these findings mean to them in the context of the programme).	Yes	

Highly Satisfactory

Executive Feedback on Section D Issues for this section relevant for feedback to senior management (positives & negatives), & justify rating. *Up to two sentences*	**Findings and conclusions are clearly presented and thoroughly discussed. However, no financial analysis is provided due to missing data and the conclusions present only the programme's strengths.**

SECTION E: RECOMMENDATIONS AND LESSONS LEARNED

Question	cc	Remarks	E/ Are the recommendations and lessons learned relevant and actionable?	Constructive feedback for future reports
Relevance and clarity of recommendations			Recommendations should be relevant and actionable to the object and purpose of the evaluation, be supported by evidence and conclusions, and be developed with involvement of relevant stakeholders. Recommendations should clearly identify the target group for each recommendation, be clearly stated with priorities for action, be actionable and reflect an understanding of the commissioning organization and potential constraints to follow up.	*Including how to address weaknesses and maintaining good practice*
43 Are the recommendations well-grounded in the evidence and conclusions reported? Recommendations should be logically based in findings and conclusions of the report.	Yes	Recommendations are based on evidence obtained from the collected data, pages 43-45. They discuss among other themes the need to scale up the project and mainstream the gender responsive approach (GRA) so that it becomes included in every teacher's training. In total there are 7 Recommendations and they relate directly to the purpose of the evaluation, i.e. assess the Relevance, Effectiveness, Efficiency, Impact and Sustainability of the programme. They are clearly presented although they could be prioritized by number.		
44 Are recommendations relevant to the object and the purpose of the evaluation? Recommendations should be relevant to the evaluated object	Yes			
45 Are recommendations clearly stated and prioritised? If the recommendations are few in number (up to 5) then this can also be considered to be prioritised. Recommendations that are over-specific or represent a long list of items are not of as much value to managers. Where there is a long list of recommendations, the most important should be ordered in priority.	Mostly			
Usefulness of recommendations			Recommendations are relevant and based on evidence obtained from the collected data. They relate directly to the purpose of the evaluation and properly identify the targeted group for action although they are not prioritised by number. Recommendations demonstrate a good understanding of the commissioning organisation (UNICEF) and awareness of the implementation constraints. The report was developed in consultation with stakeholders but it is unclear the extent to which stakeholders were involved in the development of the Recommendations. Lessons Learned are not correctly formulated and are over specific to the project.	Recommendations should be presented following a priority order. The participation of stakeholders in the formulation of the Recommendations should be clearly described. Similarly, Lessons Learned should be formulated so that the widest range of people benefit from them and so that their usefulness can transcend the particular evaluated project.
46 Does each recommendation clearly identify the target group for action? Recommendations should provide clear and relevant suggestions for action linked to the stakeholders who might put that recommendation into action. This ensures that the evaluators have a good understanding of the programme dynamics and that recommendations are realistic.	Yes	The report demonstrates good practices by identifying the targeted group for action in every given recommendation, be it UNICEF and MOEVT/PMO-RALG; DEOs, TUSEME focal persons, and teachers; or parents and community, pages 43-45. The Recommendations are realistic and appear to be potentially attainable by the responsible actors. They also demonstrate a good understanding of the commissioning organisation (UNICEF) and awareness of the implementation constraints. The report was developed in consultation with stakeholders but it is unclear the extent to which stakeholders were involved in the development of the recommendations, page 44.	Mostly Satisfactory	
47 Are the recommendations realistic in the context of the evaluation? This includes: o an understanding of the commissioning organisation o awareness of the implementation constraints o an understanding of the follow-up processes	Yes			
48 Does the report describe the process followed in developing the recommendations? The preparation of recommendations needs to suit the evaluation process. Participation by stakeholders in the development of recommendations is strongly encouraged to increase ownership and utility.	Mostly			
Appropriate lessons learned				
49 Are lessons learned correctly identified? Lessons learned are contributions to general knowledge. They may refine or add to commonly accepted understanding, but should not be merely a repetition of common knowledge. Findings and conclusions specific to the evaluated object are not lessons learned.	Mostly	Lessons learned are discussed on pages 40-41. However, they are not correctly formulated and are over-specific to the project.		
50 Are lessons learned generalised to indicate what wider relevance they may have? Correctly identified lessons learned should include an analysis of how they can be applied to contexts and situations outside of the evaluated object.	Mostly			
Executive Feedback on Section E Issues for this section relevant for feedback to senior management (positives & negatives), & justify rating. *Up to two sentences*		The Recommendations are relevant, based on collected data and properly identified the targeted group for action although they are not prioritized and stakeholder's involvement in their formulation should be better explained. Lessons Learned should be formulated so that they have a wider relevance that exceeds the context of this particular project.		

SECTION F: REPORT IS WELL STRUCTURED, LOGIC AND CLEAR

Question	cc	Remarks	F/ Overall, do all these elements come together in a well structured, logical, clear and complete report? The report should be logically structured with clarity and coherence (e.g. background and objectives are presented before findings, and findings are presented before conclusions and recommendations). It should read well and be focused.	Constructive feedback for future reports *Including how to address weaknesses and maintaining good practice*
Style and presentation				
51. Do the opening pages contain all the basic elements? Basic elements include all of: Name of the evaluated object; Timeframe of the evaluation and date of the report; Locations of the evaluated object; Names and/or organisations of evaluators; Name of the organisation commissioning the evaluation; Table of contents including tables, graphs, figures and annex; List of acronyms	Yes	The opening pages contain all of the basic elements and the report is logically structured. Appendices include appropriate elements such as the ToRs (app. F), methodological tools like TUSEME evaluation questionnaires and interview guides (app. A &B), Sampling Frame (App. C), a summary of some important findings (app. E), among others. Information that is missing from the annexes that could be useful would be the list of stakeholders interviewed. This information would strengthen the credibility of the report.		
52 Is the report logically structured? Context, purpose, methodology and findings logically structured. Findings would normally come before conclusions, recommendations & lessons learnt	Yes		The report is logically structured and presents all of the basic elements near the beginning of the document. The annexes contain appropriate and useful information that adds credibility to the report. Some important elements appear to be missing from the Executive Summary such as a description of the TUSEME project, its context as well as the purpose and scope of the evaluation. This hinders the usefulness of the Executive Summary to decision-makers and reduces its ability to stand alone.	The Executive Summary should include a summary of all of the report's key sections so as to stand alone and inform decision-making.
53 Do the annexes contain appropriate elements? Appropriate elements may include: ToRs; List of interviewees and site visits; List of documentary evidence; Details on methodology; Data collection instruments; Information about the evaluators; Copy of the evaluation matrix; Copy of the Results chain. Where they add value to the report	Yes			
54 Do the annexes increase the usefulness and credibility of the report?	Yes			
Executive Summary				
55. Is an executive summary included as part of the report? If the answer is No, question 56 to 58 should be N/A	Yes	The Executive Summary appears to be too short (pages 6-8) . Some important elements are missing, like a description of the TUSEME project, its context, information on the evaluation scope and purpose, and the evaluation methodology. This missing information hinders the usefulness of the Executive Summary to decision- makers and makes it difficult to stand alone. The writing style clearly presents a summary of Key Results and Findings, as well as a succinct overview of the recommendations.	**Mostly Satisfactory**	
56 Does the executive summary contain all the necessary elements? Necessary elements include all of: Overview of the evaluated object; Evaluation objectives and intended audience; Evaluation methodology; Most important findings and conclusions; Main recommendations	No			
57 Can the executive summary stand alone? It should not require reference to the rest of the report documents and should not introduce new information or arguments	No			
58 Can the executive summary inform decision making? It should be short (ideally 2-3 pages), and increase the utility for decision makers by highlight key priorities.	No			
Executive Feedback on Section F Issues for this section relevant for feedback to senior management (positives & negatives), & justify rating. *Up to two sentences*		**The report is overall logically structured and the annexes included provide relevant information that provides additional credibility to the report. Important information is missing from the Executive Summary (including the programme description, context, evaluation purpose and scope, and evaluation methodology), which hinders its usefulness to decision-makers.**		

Additional Information

Question	Remarks
i/ Does the evaluation successfully address the Terms of Reference? If the report does not include a TOR then a recommendation should be given to ensure that all evaluations include the TOR in the future. Some evaluations may be flawed because the TORs are inappropriate, too little time etc. Or, they may succeed despite inadequate TORs. This should be noted under vii in the next section	Yes. In general, the information found in the evaluation corresponds to the TORs. However, one major exception is in relation to the building of a counterfactual, which was called for in the ToRs but was later withdrawn from the evaluation process. A justification for why this was done is provided on page 22 as part of the Methodological Limitations.
ii/ Identify aspects of good practice in the evaluation In terms of evaluation	The description of the methodology is strong; the level of stakeholder participation, especially in terms of testing out the data collection tools, is well explained; and the recommendations are useful and realistic.
iii/ Identify aspects of good practice of the evaluation In terms of programmatic, sector specific, thematic expertise	The evaluation team demonstrated good knowledge around issues relating to the Gender Responsive Model and the Gender Responsive approach in education.

OVERALL RATING			
Question	**cc**	**Remarks**	**OVERALL RATING Informed by the answers above, apply the reasonable person test to answer the following question: Ω/ Is this a credible report that addresses the evaluation purpose and objectives based on evidence, and that can therefore be used with confidence?** This question should be considered from the perspective of UNICEF strategic management.
i/ To what extent does each of the six sections of the evaluation provide sufficient credibility to give the reasonable person confidence to act? Taken on their own, could a reasonable person have confidence in each of the five core evaluation elements separately? It is particularly important to consider: o Is the report methodologically appropriate? o Is the evidence sufficient, robust and authoritative? o Do the analysis, findings, conclusions and recommendations hold together?	Mostly	The context, the chosen methodology, conclusions, recommendations and lessons learned can be understood on their own and overall provide sufficient credibility to influence decision-makers. However, some important limitations discussed in the report including the absence of baseline data, benchmarks, and M&E information; the decision to not include a strong counterfactual methodology; missing financial information; and the inability to observe TUSEME clubs in action hinder the authority of the evaluation findings and assessments.	This is a somewhat credible report that addresses in large part the evaluation purpose and objectives based on evidence and can be used with some degree of confidence. The report observes good practices by providing a satisfactory description of the methodology used in the evaluation and provides justifications for its usage. The report is logically structured, presents well grounded findings, and useful recommendations. However, some important limitations discussed in the report including the absence of baseline data, benchmarks, and M&E information; the decision to not include a strong counterfactual methodology; missing financial information; and the inability to observe TUSEME clubs in action hinder the authority of the evaluation findings and assessments.
ii/ To what extent do the six sections hold together in a logically consistent way that provides common threads throughout the report? The report should hold together not just as individually appropriately elements, but as a consistent and logical 'whole'.	Yes	The report's structure is logical and its content helps to understand the bigger picture by providing common threads throughout the report.	Mostly Satisfactory
iii/ Are there any reasons of note that might explain the overall performance or particular aspects of this evaluation report? This is a chance to note mitigating factors and/or crucial issues apparent in the review of the report.	ToRs	The creation of a counterfactual was called for in the ToRs but was later withdrawn from the evaluation process. A justification for why this was done is provided on page 22 as part of the Methodological Limitations.	
	Other	Although a cost analysis is included in the TORs, the assessment of this is largely inconclusive as it was impossible for the evaluator to track the flow of resources. Some important limitations in the available data concerning M&E systems of the programme and information regarding its implementation hindered the robustness of the evaluation findings.	
Executive Feedback on Overall Rating Issues for this section relevant for feedback to senior management (positives & negatives), & justify rating. *Up to two sentences*	The report is logically structured, presents findings based on the collected data, and provides useful and actionable recommendations. However, some important evaluation limitations including the absence of baseline data, benchmarks, and M&E information; the decision to not include a strong counterfactual methodology; missing financial information; and the inability to observe TUSEME clubs in action hinder the authority of the evaluation findings and assessments.		