

<b>Evaluation title</b>	Final Evaluation of Disaster Preparedness and Response/Climate Change Adaptation Activities under the Office of Foreign Disaster Assistance Fund in the Philippines	<b>Evaluation Report #</b>	
<b>Type</b>	Thematic Area	<b>Centralised/ Decentralised</b>	Decentralised
<b>Global / Region or Country</b>	Philippines	<b>PHQA date</b>	26-07-2018

<b>Quality Rating - overall category</b>	<b>EPI – Overall Report Category</b>	
Meets requirements: 60% - 74%	4-7 points = Approaches requirements	
Overall, the report is of good quality. Areas for improvement include: i) A theory-based evaluation should include a programme theory or reconstruction of the theory. Changing logframes for all phases of the programme should have been a motivation to show changes in the programme theory over time. ii) conclusions and recommendations should be aligned to the evaluation criteria or main evaluation questions; iii) the boxes containing key findings should not be mixed with conclusions.		
<b>CRITERION 1: REPORT SUMMARY</b>	<b>Category</b>	<b>Meets</b>
The summary is succinct and readable. It contains the main features of the methodology and a list of evaluation users. Findings, conclusions and recommendations are well summarized and do not show any omissions. The high-quality summary shows only minimal weaknesses. Areas of improvement include the description of the evaluation subject is only partly complete, and the geographical coverage of the programme and resources raised against budget were also omitted.		
<b>CRITERION 2: OVERVIEW OF EVALUATION SUBJECT</b>	<b>Category</b>	<b>Meets</b>
The report shows and explains changes to the programme over time. A useful timeline includes key events influencing the program context. The overview section is built around a programme fact sheet. Internal programme data is presented. It is important to note that a logframe or theory of change is neither presented nor reconstructed. The ToR state that each program phase had its own logframe with different objectives. This is not further addressed in the overview section or other sections of the report. For a theory-based evaluation, this is a serious shortcoming.		
<b>CRITERION 3: EVALUATION CONTEXT, PURPOSE AND SCOPE</b>	<b>Category</b>	<b>Exceeds</b>
Overall, the quality for the context, purpose, objectives and scope is high. The learning objectives and accountability are linked to the desired holistic review of the intervention. Concerning the context, processes and actors in the shift from response planning and reactive approaches towards stronger preparedness and mitigation in the country are outlined. Limitations are minimal. These include i) The relevant geographic areas are not further specified in the context section; ii) The evaluation's objective, rather than the purposes and scope describes the balance between learning and accountability.		
<b>CRITERION 4: METHODOLOGY</b>	<b>Category</b>	<b>Meets</b>
The methodology is well presented. References are made to Annexes but the main methodological features are presented in the main report. Seven pages of methodology is rather exceptional for a main report.		

While some tables could have been annexed, it shows the methodological rigour of the evaluation. The methodology section contains a table (Table 4) providing the evidence for the alignment of the evaluation questions to the evaluation criteria. The evaluation matrix is complete and of good quality. As the report states that this is a theory-based evaluation, the reader would have expected a reconstruction of the programme theory. Changing logframes for all phases of the programme should have been a motivation to show changes in the programme theory over time. As such the ET did not succeed in implementing the theory-based evaluation approach.

**CRITERION 5: FINDINGS AND ANALYSIS**

Category

Meets

The findings section is of high quality. Results from the primary data collection are transparently presented in graphics. This contributes to the strengthening the evidence base and findings are appropriately triangulated. Table 4 shows the different data sources for each evaluation question. The only shortcoming relates to the mix of key findings and conclusions.

**CRITERION 6: CONCLUSIONS**

Category

Meets

Conclusions are based on the evidence presented in the findings section. The conclusions address both positive and negative findings. The presentation is balanced. A weakness of the conclusions is they are not listed by criterion or key evaluation questions which does not facilitate a strategic overview of the findings. The logical flow from findings to conclusions is also hampered by mixing conclusions in the key findings boxes. A clear distinction between findings and conclusions is challenging and this affects the logical flow to the conclusions section.

**CRITERION 7: GENDER**

Category

Approaches

Findings, conclusion and recommendations reflect a gender analysis. While the ET systematically explored the awareness of gender-sensitive DRRM practice, no end-beneficiary interviews took place. To the extent required, GEEW considerations were included in the data collection methods. The report assesses the inclusion of GEEW objectives in the programme design, or lack thereof. However, the report is equity-blind. This is in response to ToR that do not prioritise equity issues. The report also does not explain to what extent GEEW issues were considered for data screening and analysis. Gender indicators are qualitative and could be strengthened by providing a mix. The ToR mentions UNEG guidelines for Integrating Human Rights and Gender Equality in Evaluations in the UN System but this is not referred to in the Evaluation Report.

**CRITERION 8: RECOMMENDATIONS**

Category

Meets

The recommendations are relevant for the evaluation's objectives. The recommendations are feasible, clearly targeted, specific, at times prescriptive but very actionable after all. Constraints are mentioned where the implementation of a recommendation might be more challenging. Immediate opportunities and future programming issues are distinguished in the recommendations section. The timing for implementing recommendations is stated for each recommendation. One gap relates to the poor overall program planning with a lack of assumptions and changing logframes. This issue, despite being critical, is not fully reflected in the conclusions section and missing from the recommendations. Another shortcoming relates to recommendations not being aligned to the evaluation criteria or main evaluation questions.

**CRITERION 9: ACCESSIBILITY/CLARITY**

Category

Exceeds

The use of graphics highlighting survey results fully uses the data collected by the ET. This strengthens the evidence base and is in contracts to many other WFP evaluations reviewed. This is a good practice worth replicating across WFP. Sources are provided for all data. The good practice of analysing, systematically presenting and referencing quantitative data or the quantification of qualitative data from surveys merits special mention. The following two limitations emerge: i) Both, conclusions and recommendations are not aligned to the evaluation criteria or main evaluation questions. Strategically, this weakens the evaluation which created a strong evidence base; ii) The key findings are mixed with conclusions and cannot be distinguished as such. The logic between the sections is affected by this approach.

<b>Criteria Scoring Scale Legend - Gender Integration EPI</b>
3 points = Fully integrated
2 points = Satisfactorily integrated
1 point = Partially integrated
0 point = Not at all integrated

1. Scope & Indicators	1
2. Criteria & Questions	1
3. Methodology	1
4. Findings, Conclusions & Recommendations	2
<b>Overall EPI SCORE</b>	<b>5</b>

<b>Quality Rating Scale Legend Evaluation Reports</b>	<b>Overall Scoring of Gender EPI Scale Legend Evaluation Reports</b>
Exceeds requirements: 75% - 100%	UN SWAP Evaluation Performance Indicator
Meets requirements: 60% - 74%	11-12 points = Exceeds Requirements
Approaches requirements: 50% - 59%	8-10 points = Meets requirements
Partially meets requirements: 25% - 49%	4-7 points = Approaches requirements