Evaluation title	South Sudan: An evaluation of WFP's Portfolio (2011-16)	Evaluation Report #	OEV/2016/003
Type	Country Portfolio	Centralised/ Decentralised	Centralised
Global / Region or Country	South Sudan	PHQA date	February 2018

Quality Rating - overall category	EPI – Overall Report Category
Meets requirements: 60% - 74%	Approaches requirements= 4-7 points

The evaluation report has many strengths. It is generally robust and deals in a perceptive manner with a challenging evaluation. The report is presented in a professional manner - it uses appropriate, technically correct language, free from errors. The evaluation methodology is well thought-through and fit-for-purpose for a strategic evaluation of this nature. Anchored by a robust Evaluation Matrix that enables systematic data collection and integrated analysis to address key evaluation questions, the report presents the findings, conclusions and recommendations in a coherent manner. It also enables the investigation of GEEW and equity dimensions relevant to the evaluation subject. Drawing on a range of reliable sources, the overview of the evaluation subject and implementing context is thorough, comprehensive and insightful. It provides an informative background for understanding the evaluation findings. The evaluation findings, conclusions and recommendations are clear and address the evaluation questions and criteria in a coherent manner.

There are a number of ways in which the report could have been improved. The Executive Summary is over the maximum limit of 5,000 words, but does not include important information such as the evaluation purpose, objectives and questions. The inclusion of many figures and tables contributes to the summary's word count. Overall, GEEW and equity are dealt with in a realistic manner and the evaluation demonstrates a clear awareness of the significance of gender and equity dimensions in the portfolio. However, the focus on gender in the context overview, analysis, conclusions and recommendations could have been strengthened, and where gender-disaggregated data is not available, this should be have been pointed out and identified as a limitation of the evaluation. A (more comprehensive) summary of the evaluation methodology could also have explained how gender-sensitive data collection was conducted. The report could have provided a short overview of specific lessons, challenges or recommendations for conducting gender-responsive strategic evaluations of this nature. The inter-relationship between recommendations could have been made more explicit. Readability of the report could be enhanced by including concise, high-level summaries of essential information from the Inception Report and Annexes in the main report; however, this would require a trade-off with the page limit. Readability could also have been enhanced by using bold text (or similar) to highlight key messages.

CRITERION 1: REPORT SUMMARY	Category	Approaches

The Summary provides a clear, succinct overview of the complex evaluation subject. Key evaluation findings and conclusions (with some exceptions) are summarised in a perceptive manner, and all recommendations are summarised. The Executive Summary is over the maximum limit of 5,000 words, but does not include important information such as the evaluation purpose, objectives and questions. The inclusion of many

figures and tables contributes to the summary's word count. Only the most relevant tables (summarised as appropriate), i.e. that directly support of the narrative should be included in the summary.

CRITERION 2: OVERVIEW OF EVALUATION SUBJECT	Category	Exceeds

The overview of the evaluation subject is thorough, comprehensive and insightful. The history, dynamics and evolution of WFP's involvement in South Sudan are dealt with, and the rationale for operations that are within the scope of the evaluation is clarified in relation to contextual challenges and changes. The overview of WFP's engagement and specific operations in South Sudan does not address gender issues however. There is also no reference to lessons from previous evaluations and reviews, and how they informed and shaped the portfolio of operations that Is subject to evaluation.

CRITERION 3: EVALUATION CONTEXT, PURPOSE AND SCOPE	Category	Exceeds

Drawing on a range of reliable sources, the overview of the implementing context in South Sudan is comprehensive, relevant and insightful to the evaluation subject. It provides a highly informative background for understanding the evaluation findings. However, the context overview does not deal systematically with gender as a cross-cutting issue. Where gender-disaggregated data is not available, this should have been pointed out and identified as a limitation of the evaluation. Furthermore, the evaluation purpose, objectives and scope should have been dealt with systematically and in more detail.

CRITERION 4: METHODOLOGY	Category	Meets

Given the wide-ranging scope of the evaluation and challenging implementing context, the evaluation is based on a well-thought-through and fit-for-purpose methodology. There is coherence between the evaluation purpose, objectives, criteria (with the exception of coverage) and questions. The Evaluation Matrix is detailed and provides assurance that evaluation questions are addressed in a systematic manner, based on evidence from multiple sources. However, although minimum information on the methodology is provided in the Inception Report and Annexes of the Evaluation Report, and these are referenced in the main report, it would have been useful to include short summaries of key elements of the methodology in the Evaluation Report, e.g. evaluation criteria and questions (including clarification around the exclusion of coverage as an evaluation criterion); evaluability challenges and mitigation measures, etc. The methodology section in the Evaluation Report itself does not provide an adequate overview of the methodology.

CRITERION 5: FINDINGS AND ANALYSIS	Category	Exceeds

Findings provide a comprehensive, coherent and balanced overview of the program's performance against the key evaluation questions. It deals perceptively with complex issues in a highly challenging implementing context, highlighting and explaining the effect that this has had on the programme. The section on findings is robust and there are no obvious weaknesses that stand out. If a weakness has to be singled out, it would be that findings dealing with follow-up to previous evaluations, e.g. the Inter-Agency Humanitarian Evaluation (IAHE) and WFP gender policy evaluation are not made explicit.

CRITERION 6: CONCLUSIONS	Category	Meets

The conclusions are based on a perceptive analysis of a comprehensive evidence base and it incorporates all salient findings in a manner that address both the key evaluation questions and evaluation criteria. The conclusions are generally robust and no major weaknesses stand out. However, connectedness is raised for the first time in the conclusions so it does not flow logically from prior findings and analysis.

CRITERION 7: GENDER	Category	Meets

The evaluation methodology, analysis, findings, conclusions and recommendations demonstrate an awareness and realistic efforts to address gender and equity dimensions in a manner that is relevant for a strategic evaluation of this nature. A comprehensive Evaluation Matrix that includes general and specific questions about GEEW and equity ensures that it remains in the evaluation's line of sight throughout. The methodology does not include guidance to ensure that data collection (interviews with beneficiaries, document reviews and key informant interview) are conducted in a gender-sensitive manner, however,

Furthermore, GEEW is not dealt with explicitly in the evaluation recommendations and the report offers no specific lessons, challenges or recommendations for conducting gender-responsive strategic evaluations of this nature.

CRITERION 8: RECOMMENDATIONS	Category	Meets

The recommendations, which flow logically from the findings and recommendations, are aligned with the purpose and expectation of the evaluation. Recommendations are specific, time-bound and responsibility for following up and implementing recommendations are assigned to specific stakeholders. Although the recommendations are generally actionable, the successful implementation of some recommendations (e.g. maximising humanitarian-development synergies and the development of various strategies to strengthen and guide a multitude of operations) would depend on the successful implementation of others (e.g. addressing staffing of the Country Office and resource mobilisation to fund a strategic portfolio). This could be made more explicit.

CRITERION 9: ACCESSIBILITY/CLARITY	Category	Meets

Although slightly over the 50-page limit, the report is easy to read. It is presented in a professional manner - it uses appropriate, technically correct language, free from errors. The report presents a balanced, credible assessment of a large country portfolio and the audiences for which it is intended should find it useful and accessible. The presentation could be strengthened by using bold text or other means to highlight key messages. Readability of the report could be enhanced by including concise, high-level summaries of essential information from the Inception Report and Annexes in the main report; however, this will require a trade-off with the page limit. Remaining typing and punctuation errors, as well as missing references, should be addressed.

Criteria Scoring Scale Legend - Gender Integration EPI		1. Scope & Indicators	2
3 points = Fully integrated		2. Criteria & Questions	2
2 points = Satisfactorily integrated		3. Methodology	2
1 point = Partially integrated		4. Findings, Conclusions & Recommendations	1
0 point = Not at all integrated		Overall EPI SCORE	7

Quality Rating Scale Legend Evaluation Reports		Overall Scoring of Gender EPI Scale Legend Evaluation Reports
		UN SWAP Evaluation Performance Indicator
Exceeds requirements:	75% - 100%	
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
Does not meet requirements:	0% - 24%	0-3 points = Missing requirements