

Organizational unit: UNFPA Country Office Papua New Guinea **Year of report:** 2016

Title of evaluation report: UNFPA Country Program Evaluation: Papua New Guinea, 2012-2017

Overall quality of report: **Very good** **Date of assessment:** 15 Dec. 16

Overall comments: The evaluation report is structured in line with UNFPA standards. The executive summary is stand-alone section and is drafted in a professional manner. The evaluation is remarkably clear on its purpose and structure. The evaluation design and methodology are appropriate and are clearly outlined in the methodology section. Available data sources were effectively used and the quality of data is ensured through mixed methods and triangulation. Data analysis is gender disaggregated and reflects contextual factors. Findings are concise and well substantiated by evidence. Conclusions are linked to findings and are organized by evaluation questions. Though they are strategic they are not always sufficiently specific. Recommendations flow logically from conclusions, are very specific and well prioritized, facilitating an appropriate management response. While gender components were discussed under findings, conclusions and recommendations, no specific gender analysis was conducted.

Assessment Levels

Very good:	strong, above average, best practice	Good:	satisfactory, respectable	Fair:	with some weaknesses, still acceptable	Unsatisfactory:	weak, does not meet minimal quality standards
-------------------	--------------------------------------	--------------	---------------------------	--------------	--	------------------------	---

Quality Assessment Criteria	<i>Insert <u>assessment level</u> followed by main <u>comments</u>. (use 'shading' function to give cells corresponding colour)</i>	
<p>1. Structure and Clarity of Reporting</p> <p>To ensure the report is comprehensive and user-friendly</p> <ul style="list-style-type: none"> Is the report easy to read and understand (i.e. written in an accessible non-technical language appropriate for the intended audience)? Is the report focused and to the point (e.g. not too lengthy)? Is the report structured in a logical way? Is there a clear distinction made between analysis/findings, conclusions, recommendations and lessons learned (where applicable)? Do the annexes contain – at a minimum – the ToRs; a bibliography, a list of interviewees, the evaluation matrix and methodological tools used (e.g. interview guides; focus group notes, outline of surveys)? <p>Executive summary</p> <ul style="list-style-type: none"> Is an executive summary included in the report, written as a stand-alone section and presenting the main results of the evaluation? Is there a clear structure of the executive summary, (i.e. i) Purpose, including intended audience(s); ii) Objectives and brief description of intervention; iii) Methodology; iv) Main 	Assessment Level:	Very good
	<p>Comment:</p> <p>The report has the expected structure, including all of the expected annexes. It included a results-matrix that was agreed with the review committee. The report is written in plain professional language. It is different from other UNFPA evaluation reports in terms of clarity of findings. All findings are assigned a number, description is not too lengthy. Assumptions are numbered and each assumption is assessed individually after presenting the findings.</p> <p>The annexes contain the ToRs (p. 73), documents consulted (p. 96), a list of people interviewed (p. 93), the evaluation matrix (p. 100). There are no interview guides, focus group notes or outline of surveys among the annexes.</p> <p>The executive summary covers all of the necessary ground and by focusing on the purpose of the evaluation (to influence the next country programme) provides an excellent summary of the evaluation. This part is rated very good.</p> <p>Executive summary presents the main results of the evaluation. It has a clear structure: purpose and audience of the country program evaluation (p. 1), objectives of the evaluation and a description of the country program (p. 1), evaluation methodology (p. 2), main conclusions (p. 2), and recommendations (p. 2). Length of the executive summary is 3 pages.</p>	

<p>conclusions; v) Recommendations)?</p> <ul style="list-style-type: none"> Is the executive summary reasonably concise (e.g. with a maximum length of 5-10 pages)? 	
--	--

<p>2. Design and Methodology</p> <p><i>To ensure that the evaluation is put within its context</i></p> <ul style="list-style-type: none"> Does the evaluation describe whether the evaluation is for accountability and/or learning purposes? Does the evaluation describe the target audience for the evaluation? Is the development and institutional context of the evaluation clearly described? Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change? Does the evaluation explain any constraints and/or general limitations? <p><i>To ensure a rigorous design and methodology</i></p> <ul style="list-style-type: none"> Is the evaluation approach and framework clearly described? Does it establish the evaluation questions, assumptions, indicators, data sources and methods for 	<table border="1"> <tr> <td data-bbox="991 430 1633 521">Assessment Level:</td> <td data-bbox="1633 430 1858 521">Good</td> </tr> <tr> <td colspan="2" data-bbox="991 521 1858 1308"> <p>Comment:</p> <p>The evaluation is for accountability and learning purposes “The purpose of the Country Program Evaluation (CPE) is twofold: 1) a learning tool...; and 2) an accountability tool...” (p. 4). The evaluation is remarkably clear on its purpose and the structure. As a “One-UN” country PNG has a complex management structure in which UNFPA functions and this is well-described.</p> <p>The text clearly shows both the who and why of the evaluation “the main users of the evaluation results will be decision-makers of UNFPA at country office, regional and global level, and the organization’s Executive Board...” (p. 4).</p> <p>The political, economic and social context of PNG is well-described in the chapter 2 (section 2.1 “Development challenges and national strategies” and section 2.2 “The role of external assistance in PNG”) including particularly the areas of interest to UNFPA.</p> <p>The evaluators, as part of the design process, created the evaluation matrix (annex 4) to guide the evaluation and had it approved by the</p> </td> </tr> </table>	Assessment Level:	Good	<p>Comment:</p> <p>The evaluation is for accountability and learning purposes “The purpose of the Country Program Evaluation (CPE) is twofold: 1) a learning tool...; and 2) an accountability tool...” (p. 4). The evaluation is remarkably clear on its purpose and the structure. As a “One-UN” country PNG has a complex management structure in which UNFPA functions and this is well-described.</p> <p>The text clearly shows both the who and why of the evaluation “the main users of the evaluation results will be decision-makers of UNFPA at country office, regional and global level, and the organization’s Executive Board...” (p. 4).</p> <p>The political, economic and social context of PNG is well-described in the chapter 2 (section 2.1 “Development challenges and national strategies” and section 2.2 “The role of external assistance in PNG”) including particularly the areas of interest to UNFPA.</p> <p>The evaluators, as part of the design process, created the evaluation matrix (annex 4) to guide the evaluation and had it approved by the</p>	
Assessment Level:	Good				
<p>Comment:</p> <p>The evaluation is for accountability and learning purposes “The purpose of the Country Program Evaluation (CPE) is twofold: 1) a learning tool...; and 2) an accountability tool...” (p. 4). The evaluation is remarkably clear on its purpose and the structure. As a “One-UN” country PNG has a complex management structure in which UNFPA functions and this is well-described.</p> <p>The text clearly shows both the who and why of the evaluation “the main users of the evaluation results will be decision-makers of UNFPA at country office, regional and global level, and the organization’s Executive Board...” (p. 4).</p> <p>The political, economic and social context of PNG is well-described in the chapter 2 (section 2.1 “Development challenges and national strategies” and section 2.2 “The role of external assistance in PNG”) including particularly the areas of interest to UNFPA.</p> <p>The evaluators, as part of the design process, created the evaluation matrix (annex 4) to guide the evaluation and had it approved by the</p>					

<p>data collection?</p> <ul style="list-style-type: none"> ● Were the methods chosen appropriate for addressing the evaluation questions? Are the tools for data collection described and justified? ● Is the methods for analysis clearly described? ● Are methodological limitations acknowledged and their impact on the evaluation described? (Does it discuss how any bias has been overcome?) ● Is the sampling strategy described? Does the design include validation techniques? ● Is there evidence of involvement of stakeholders in the evaluation design? (Is there a comprehensive/credible stakeholder map?) ● Does the methodology enable the collection and analysis of disaggregated data? ● Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)? 	<p>oversight committee. The evaluation matrix contains assumptions to be assessed, indicators, sources of information, and methods for data collection. The methods chosen are appropriate for addressing the evaluation questions. For instance, to assess the assumption “Comprehensive, gender-sensitive, high-quality SRH services are in place and accessible in underserved areas with a focus on vulnerable groups in project areas» the evaluators used such methods as Document review, UNFPA CO team presentation, and others (p. 101).</p> <p>The tools for data collection are described in the methodology section and the evaluation matrix, but the evaluators do not provide detailed justification. For instance, they misplace sources of information and methods and tools for data collection in the evaluation matrix. Thus, UNFPA and UNCT partners are mentioned in the Sources, whereas the Methods of data collection include “KII DoH, DoE, RPNGC, DCD, implementers” (p. 107). The evaluators need to be clear what they call sources of data such as “Field visit to relevant institutions” and what they call methods for data collection like Documents Review.</p> <p>Results of a Theory of change analysis presented in the chapter 4.2.1. “General effectiveness findings” (findings 1-3). Re-constructed theory of change is described on the Figure 2 «Proposed intervention logic for the UNFPA PNG Country Program» and Figure 3 «Theoretical framework for the UNFPA program and interventions.» . Also, it is a mistake on the page 6 “Details of the Theory of Change and interventions logic are provided in chapter 3.1, under ‘effectiveness’ (page 31)”, whereas it is chapter 4.1 on the page 31.</p> <p>Sampling approach is well described “Sampling of field sites to visit and</p>
---	--

informants to interview was purposive...” (p. 6). The limitations of the methodology are clearly and candidly described on the pages 7-8.

The evaluation approach and framework are explained in the methodology section, including evaluation criteria (pp. 5-7, pp. 8-9). The methodology consisted of document review (from which considerable quantitative information was obtained), semi-structured interviews with a stratified sample of key stakeholders (purposive but checked against other sources), a selection of focus groups and site visits to all of the four provinces in which the UNFPA program worked.

Stakeholder involvement was built into the methodology at each of the phases: “a validation workshop with stakeholders confirmed validity and completeness of the field findings and documentation review (p. 2), “Limited access to staff was compensated by additional interviews with stakeholders and documentation review” (p. 8). Preparatory Phase included stakeholder mapping (p. 8), but annexes do not contain a stakeholder map.

Gender was a critical program area and was well-covered. The ToR required the evaluators to explain how “Approach will integrate both gender and human rights perspectives quantitative and qualitative data” (pp. 81-82). The evaluators refer to “UNEG guidance for including gender, equity and human rights...” and say that they «selected gender responsive evaluation questions, methods, tools and data analysis techniques where appropriate, and the CPE team included a gender & development expert” (p. 5).

<p>3. Reliability of Data</p>	<p>Assessment Level:</p>	<p>Good</p>
<p><i>To ensure quality of data and robust data collection processes</i></p> <ul style="list-style-type: none"> ● Did the evaluation triangulate all data collected? ● Did the evaluation make explicit any possible issues (bias, data gaps etc.) in primary and secondary data sources and if relevant, explained what was done to minimize such issues? I.e. did the evaluation make explicit possible limitations of the data collected? ● Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations? ● Is there adequate gender disaggregation of data? And if this has not been possible, is it explained? ● Does the evaluation make explicit the level of involvement of different stakeholders in the different phases of the evaluation process? 	<p>Comment:</p> <p>The data were well-triangulated and were used according to reliability: “The methodology is in accordance with UNFPA guidance, and includes mixed methods with triangulation and validation of the findings” (p. 2); “Document review and stakeholder interviews did not reveal any materially relevant unintended consequences” (p. 25).</p> <p>The evaluation team made effective use of available data sources. For instance, the evaluation team accessed only two evaluation reports, because other UNFPA program documents did not provide evidence of effectiveness (p. 34). But, the analysis lack of references in some cases, for instance, Effectiveness finding # 5 states that “UNFPA has contributed to significant progress in national policy development, although the specific contribution of UNFPA to the joint UN policy dialogue is hard to identify.” The evaluators do not justify “significant progress” and “hard to identify”.</p> <p>Methodology includes very detailed description of the limitations (p. 5, p. 7). Also, the evaluators explain limitations in the analysis, for instance, they write “UNFPA guidance on CPE recognizes the limitations of CCPD result frameworks...” (p. 32). The evaluators were aware of possible representativeness issues based on the purposive sampling, but tested the reliability of the data through a meeting with stakeholders prior to the evaluator’s departure from the field mission part.</p> <p>The evaluation consultants collected gender disaggregated data that is</p>	

	<p>evident from the text and from the annex 2A “People interviewed.” They have made an effort to keep involvement of different stakeholders clear: “Limited access to staff was compensated by additional interviews with stakeholders and documentation review” (p. 8).</p>
--	--

<p>4. Analysis and Findings</p> <p><i>To ensure sound analysis</i></p> <ul style="list-style-type: none"> ● Is information analysed and interpreted systematically and logically? ● Are the interpretations based on carefully described assumptions? ● Is the analysis presented against the evaluation questions? ● Is the analysis transparent about the sources and quality of data? ● Are possible cause and effect links between an intervention and its end results explained? ● Where possible, is the analysis disaggregated to show different outcomes between different target groups? ● Are unintended results identified? 	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 80%; text-align: right;">Assessment Level:</td> <td style="background-color: #ADD8E6; text-align: center;">Very good</td> </tr> <tr> <td colspan="2"> <p>Comment:</p> <p>The evaluators present logical analysis based on documents review and interviews with stakeholders, therefore, findings are substantiated by evidence. The analysis starts from the Theory of Change re-construction, review of the 5th UNFPA country program (2012-2015), and answers to the ten evaluation questions, including assumptions.</p> <p>All findings are assigned a number and supported by discussion of evidence. The analysis explains possible cause and effect links between outputs, outcomes and impacts, for instance, it is said that “UNFPA enabled roll out of the MSI FP training and the RHTU obstetric care training, but the coverage remains too low for impact. UNFPA supported training on supply, security and stock management for health workers, resulting in reduced stock outs” (p. 35).</p> <p>Where possible, the analysis is disaggregated to show different outcomes between target groups, for instance, evidence shows that “There is anecdotal evidence that implants are extremely popular</p> </td> </tr> </table>	Assessment Level:	Very good	<p>Comment:</p> <p>The evaluators present logical analysis based on documents review and interviews with stakeholders, therefore, findings are substantiated by evidence. The analysis starts from the Theory of Change re-construction, review of the 5th UNFPA country program (2012-2015), and answers to the ten evaluation questions, including assumptions.</p> <p>All findings are assigned a number and supported by discussion of evidence. The analysis explains possible cause and effect links between outputs, outcomes and impacts, for instance, it is said that “UNFPA enabled roll out of the MSI FP training and the RHTU obstetric care training, but the coverage remains too low for impact. UNFPA supported training on supply, security and stock management for health workers, resulting in reduced stock outs” (p. 35).</p> <p>Where possible, the analysis is disaggregated to show different outcomes between target groups, for instance, evidence shows that “There is anecdotal evidence that implants are extremely popular</p>	
Assessment Level:	Very good				
<p>Comment:</p> <p>The evaluators present logical analysis based on documents review and interviews with stakeholders, therefore, findings are substantiated by evidence. The analysis starts from the Theory of Change re-construction, review of the 5th UNFPA country program (2012-2015), and answers to the ten evaluation questions, including assumptions.</p> <p>All findings are assigned a number and supported by discussion of evidence. The analysis explains possible cause and effect links between outputs, outcomes and impacts, for instance, it is said that “UNFPA enabled roll out of the MSI FP training and the RHTU obstetric care training, but the coverage remains too low for impact. UNFPA supported training on supply, security and stock management for health workers, resulting in reduced stock outs” (p. 35).</p> <p>Where possible, the analysis is disaggregated to show different outcomes between target groups, for instance, evidence shows that “There is anecdotal evidence that implants are extremely popular</p>					

- Is the analysis presented against contextual factors?
- Does the analysis include reflection of the views of different stakeholders (reflecting diverse interests)? E.g. how were possible divergent opinions treated in the analysis?
- Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?

To ensure credible findings

- Can evidence be traced through the analysis into findings? E.g. are the findings substantiated by evidence?
- Do findings follow logically from the analysis?
- Is the analysis of cross-cutting issues integrated in the findings?

among women, but also about severe resistance among male partners due to misconceptions” (p. 35).

The analysis is presented against contextual factors (chapter 2. Country context). The analysis examines all possible results as the “evaluation found no evidence that UNFPA program activities, especially those subcontracted to implementing partners, are based on project documents including situation analysis and rationale...” (p. 27). The analysis reflects the views of different stakeholders (Annex 2. List of persons / Institutions met, Annex 4. Evaluation matrix).

The analysis includes findings on cross-cutting issues in different parts of the report, for instance, there is detailed discussion about vulnerable groups “Some of the interventions supported by UNFPA through implementing partners do not seem to address the most vulnerable subgroups...” (p. 28).

5. Conclusions	Assessment Level:	Good
<p><i>To assess the validity of conclusions</i></p> <ul style="list-style-type: none"> • Are conclusions credible and clearly related to the findings? • Are the conclusions demonstrating an appropriate level of analytical abstraction? • Are conclusions conveying the evaluators' unbiased judgement of the intervention? 	<p>Comment:</p> <p>Conclusions chapter is well structured and logical. Conclusions are related to the findings. The consultants connect each conclusion to the Evaluation criteria and Evaluation questions. They provide justification for all conclusions.</p> <p>Presented conclusions of the CPE are all strategic. There are no programmatic conclusions. The evaluators comment on such decision, "issues regarding implementation of individual interventions are discussed in chapter 4 and adequately reflected in Chapter 6 (Recommendations)" (p. 63).</p> <p>Some conclusions need references to the findings to sound reliable, for instance, it is said that "The supportive policy environment for population, development and SRHR is sustainable in the medium term" (p. 64). It is unclear how the consultants come up with this conclusion. In general, references to the findings would add credibility to all conclusions.</p> <p>Some conclusions are not specific, for instance, the consultants state that "UNFPA is in a good position to support national and provincial health authorities to address these barriers, and capitalize on the recent policy advances" (p. 64). Clarification of the phrase "good position" would be useful.</p>	

<p>6. Recommendations</p> <p><i>To ensure the usefulness and clarity of recommendations</i></p> <ul style="list-style-type: none"> ● Do recommendations flow logically from conclusions? ● Are the recommendations sufficiently clear, targeted at the intended users and operationally-feasible? ● Do recommendations reflect stakeholders’ consultations whilst remaining balanced and impartial? ● Is the number of recommendations manageable? ● Are the recommendations prioritised and clearly presented to facilitate appropriate management response and follow up on each specific recommendation? 	<p>Assessment Level:</p>	<p>Very good</p>
<p>Comment:</p> <p>There are 4 recommendations that are based on strategic conclusions. The recommendations are prioritised (High or Medium) and clearly described to facilitate appropriate management response and follow up on each specific recommendation. Three recommendations are ranked “High”. One recommendation has Medium level.</p> <p>All recommendations are assigned a Target Level. Most of them are directed at the UNFPA country program management office. The recommendations sound specific, for instance, Recommendation 1 states that Country office should “Ensure that the design of the country program and individual interventions are based on <i>formative research and international good practice</i>; develop and use a M&E systems as per corporate guidance, and undertake operational research to document lessons and inform policy dialogue” (p. 68).</p> <p>Recommendations are operationally-feasible. They contain operational implications with clear action points, for instance, the consultants advise to complete a SWOT analysis of “provincial health systems to plan and deliver family planning services” and provide “assessment of abortion and post-abortion services in PNG” (p. 68).</p> <p>It is nothing said in the chapter if the recommendations reflected stakeholders’ consultations. The Methodology explains that “The field phase ended with an internal data analysis meeting, validation</p>		

	workshop... The evaluation team leader drafted the evaluation report, taking into account comments made at the validation meeting/debrief” (pp. 8-9).
--	---

<p>7. Gender</p> <p><i>To assess the integration of Gender Equality and Empowerment of Women (GEEW)¹</i></p> <ul style="list-style-type: none"> ● Is GEEW integrated in the evaluation scope of analysis and indicators designed in a way that ensures GEEW-related data to be collected? ● Do evaluation criteria and evaluation questions specifically address how GEEW has been integrated into design, planning, implementation of the intervention and the results achieved? ● Have gender-responsive evaluation methodology, methods and tools, and data analysis techniques been selected? ● Do the evaluation findings, conclusions and recommendations reflect a gender analysis? 	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="text-align: right; padding-right: 10px;">Assessment Level:</td> <td style="background-color: #FFD700; text-align: center; padding: 5px;">Fair</td> </tr> <tr> <td colspan="2" style="padding: 10px;"> <p>Comment:</p> <p>GEEW is explicitly mentioned as a cross-cutting theme to be addressed within the scope of the evaluation.</p> <p>Indicators are designed in a way that ensures some gender-related data to be collected, for instance, “Uptake of EOC & FP services increased, especially for vulnerable women and men” (p. 101), “National (& provincial) Womens’ Councils functional” (p. 107). However, they do not allow for a full-fledged assessment of GEEW integration.</p> <p>The report states that «The CPE selected gender responsive evaluation questions, methods, tools and data analysis techniques where appropriate, and the CPE team included a gender & development expert” (p. 5). However, there is no evidence of how GEEW was indeed factored in data collection tools, such as interview protocols.</p> <p>The evaluators claim that “Quantitative data on programs and services was disaggregated for sex and age where possible and appropriate”</p> </td> </tr> </table>	Assessment Level:	Fair	<p>Comment:</p> <p>GEEW is explicitly mentioned as a cross-cutting theme to be addressed within the scope of the evaluation.</p> <p>Indicators are designed in a way that ensures some gender-related data to be collected, for instance, “Uptake of EOC & FP services increased, especially for vulnerable women and men” (p. 101), “National (& provincial) Womens’ Councils functional” (p. 107). However, they do not allow for a full-fledged assessment of GEEW integration.</p> <p>The report states that «The CPE selected gender responsive evaluation questions, methods, tools and data analysis techniques where appropriate, and the CPE team included a gender & development expert” (p. 5). However, there is no evidence of how GEEW was indeed factored in data collection tools, such as interview protocols.</p> <p>The evaluators claim that “Quantitative data on programs and services was disaggregated for sex and age where possible and appropriate”</p>	
Assessment Level:	Fair				
<p>Comment:</p> <p>GEEW is explicitly mentioned as a cross-cutting theme to be addressed within the scope of the evaluation.</p> <p>Indicators are designed in a way that ensures some gender-related data to be collected, for instance, “Uptake of EOC & FP services increased, especially for vulnerable women and men” (p. 101), “National (& provincial) Womens’ Councils functional” (p. 107). However, they do not allow for a full-fledged assessment of GEEW integration.</p> <p>The report states that «The CPE selected gender responsive evaluation questions, methods, tools and data analysis techniques where appropriate, and the CPE team included a gender & development expert” (p. 5). However, there is no evidence of how GEEW was indeed factored in data collection tools, such as interview protocols.</p> <p>The evaluators claim that “Quantitative data on programs and services was disaggregated for sex and age where possible and appropriate”</p>					

¹ This assessment criteria is fully based on the UN-SWAP Scoring Tool, see Annex 7. Each sub-criteria shall be equally weighted (in correlation with the calculation in the tool and totalling the scores 11-12 = very good, 8-10 = good, 4-7 = Fair, 0-3=unsatisfactory). One question is if this criteria should be included in the overall evaluation quality assessment grid, or form a separate column and be assessed on its own.

	<p>(p. 5). It is possible to find examples of data-disaggregation in the Country context section, for instance, “over two thirds of unmarried males and females aged 15–24 have had sex, with 16 as the median age of first sex” (p. 13), “Women and girls have substantially less access to health care and education services than males” (p. 14).</p> <p>Gender issues are well described in all sections of the Findings chapter. They are also discussed in the conclusions and recommendations, although no specific gender analysis has been conducted.</p>
--	--

Overall Evaluation Quality Assessment

	Assessment Levels (*)			
Quality assessment criteria (scoring points*)	Very good	Good	Fair	Unsatisfactory
1. Structure and clarity of reporting, including executive summary (7)	7			
2. Design and methodology (13)		13		
3. Reliability of data (11)		11		
4. Analysis and findings (40)	40			
5. Conclusions (11)		11		
6. Recommendations (11)	11			
7. Integration of gender (7)			7	
Total scoring points	58	35	7	
Overall assessment level of evaluation report	Very good			
	Very good • very confident to use	Good • confident to use	Fair • use with caution	Unsatisfactory • not confident to use

(*) (a) Insert scoring points associated with criteria in corresponding column (e.g. - if 'finding and analysis' has been assessed as 'good', enter 40 into 'Good' column). (b) Assessment level with highest 'total scoring points' determines 'Overall assessment level of evaluation report'. Write corresponding assessment level in cell (e.g. 'Fair'). (c) Use 'shading' function to give cells corresponding colour.

If the overall assessment is 'Fair', please explain²:

- How it can be used?
- What aspects to be cautious about?

Where relevant, please explain the overall assessment Very good, Good or Unsatisfactory³:

This is an evaluation that systematically produces good data, from various sources, and combines them into carefully presented findings, conclusions and recommendations. It is a model.

Consideration of significant constraints⁴

The quality of this evaluation report has been hampered by exceptionally difficult circumstances:

yes x no

If yes, please explain:

² The purpose here is to clarify in what way the report can be used. This in order to assist the elaboration of a relevant Management Response and the wider use of the evaluation findings back into programming. When a report has been assessed as Fair, it is obligatory to fill this text box in.

³ The purpose is, where relevant, to clarify for example severe unbalances in the report (for example, the report is good overall but recommendations very weak). Is optional to fill in.

⁴ E.g. this should only be used in case of significant events that has severely hampering the evaluation process like natural disasters, evaluators falling sick, unexpected significant travel restrictions, etc. More 'normal' limitations should be mentioned under relevant section above.