



|                                    |  |                            |              |
|------------------------------------|--|----------------------------|--------------|
| <b>Organizational unit:</b>        | Evaluation Office  | <b>Year of report:</b>     | 2017         |
| <b>Title of evaluation report:</b> | Evaluation of the Architecture Supporting the Operationalisation of the UNFPA Strategic Plan 2014-17   |                            |              |
| <b>Overall quality of report:</b>  | <b>Very Good</b>   | <b>Date of assessment:</b> | 31 July 2017 |
| <b>Overall comments:</b>           | <p>The evaluation report is written in line with UNFPA standards. The factors which make the solid report include the clarity and conciseness of the text, consistency in the reporting structure, and that it was easy to read and understand. What was particularly notable in this evaluation was its explanation of a credible stakeholder consultation process. The executive summary is a stand-alone section and is drafted in a professional manner. The evaluation design and methodology are appropriate and are clearly explained in the methodology section. Available data sources were effectively utilized and the quality of data is ensured through triangulation. Data analysis reflects contextual factors.</p> <p>Findings are concise and well substantiated by evidence.</p> <p>Conclusions are linked to findings and are organized by evaluation questions. Conclusions focus at the macro level on three strategic key messages that UNFPA management needs to address in improving the support architecture. Recommendations flow logically from conclusions and are in fact cross-referenced at the beginning of this section. Recommendations are clear and targeted to the appropriate unit of UNFPA; whether they are operationally feasible is not immediately self-evident given they are formulated in a manner best understood by internal staff, although the management response suggests that they are feasible.</p> <p>While gender issues were discussed under findings, conclusions and recommendations, they do not reflect a gender analysis. The evaluators do not explain how they applied UNEG guidance on gender equality and human rights, even though they indicate that they followed them.</p> <p>The voices of both men and women are reflected as it is indicated in the list of persons interviewed and consulted, but not all the people are disaggregated in terms of gender.</p> |                            |              |

|                          |   |                                       |  |   |
|--------------------------|---|---------------------------------------|--|---|
| <b>Assessment Levels</b> | <b>Very Good</b> strong, above average, best practice | <b>Good</b> satisfactory, respectable | <b>Fair</b> with some weaknesses, still acceptable | <b>Unsatisfactory</b> weak, does not meet minimal quality standards |
|--------------------------|---|---------------------------------------|--|---|

|  |                      |  |
|--|----------------------|--|
| <b>Quality Assessment Criteria</b>   |                      |  |
| <b>I. Structure and Clarity of Reporting</b>   | Yes<br>No<br>Partial | Assessment Level: <b>Very good</b>   |
| <i>To ensure the report is comprehensive and user-friendly</i>   |                      | Comment:   |
| <b>I. Is the report easy to read and understand (i.e. written in an accessible language appropriate for the intended audience) with minimal grammatical, spelling or punctuation errors?</b> | Yes                  | <b>The evaluation report is written in line with UNFPA standards.</b> The evaluation has clear structure, presented in an easy to read and understandable, non-technical manner. It is well organized and thorough, focused and to the point. The annexes include the required items.<br>The report is reasonable in length, being 53 pages. |

|  |     |  |
|--|-----|--|
| 2. Is the report of a reasonable length? (maximum pages for the main report, excluding annexes: 60 for institutional evaluations; 70 for CPEs; 80 for thematic evaluations)  | Yes | <p>The executive summary is stand-alone section and is drafted in a professional and clearly structured manner. The Executive Summary covers the required items. It is concise. There are two versions of the Executive Summary: the one in the body of the text and a separate self-standing one. The later is the shorter 3 page version and the former is the longer 4 page version. The self-standing short version includes "displays" aka boxes and graphs which include, inter alia, the evaluation questions cross-referenced to the major findings, which is a very effective technique. It follows a parallel structure to the main body of the report and functions as a stand-alone section. It distinguishes the "primary purpose" of the evaluation from the three core objectives which focus on macro-level issues.</p> <p>The minor issue is use of several terms (e.g., the bulls-eye, the pink quadrant, and the colored coding of countries) all of which could have been explained in a footnote to help the outside reader understand what is meant.</p> |
| <b>3. Is the report structured in a logical way? Is there a clear distinction made between analysis/findings, conclusions, recommendations and lessons learned (where applicable)?</b>   | Yes |  |
| 4. Do the annexes contain – at a minimum – the ToRs; a bibliography; a list of interviewees; the evaluation matrix; methodological tools used (e.g. interview guides; focus group notes, outline of surveys) as well as information on the stakeholder consultation process? | Yes |  |
| <i>Executive summary</i>   |     |  |
| <b>5. Is an executive summary included in the report, written as a stand-alone section and presenting the main results of the evaluation?</b>  | Yes |  |
| 6. 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 conclusions; v) Recommendations)?   | Yes |  |
| 7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?  | Yes |  |

|  |                      |  |      |
|--|----------------------|--|------|
| 2. Design and Methodology  | Yes<br>No<br>Partial | Assessment Level:  | Good |
| <i>To ensure that the evaluation is put within its context</i>   |                      | Comment: <b>The evaluation design and methodology are appropriate and are clearly explained in the methodology section.</b>  |      |
| 1. Does the evaluation describe the target audience for the evaluation?  | Yes                  | While the chapter does not explicitly "describe the target audience" by definition, the two target audiences (the Executive Board and internal Management) are implicit.   |      |
| 2. Is the development and institutional context of the evaluation clearly described and constraints explained?   | Yes                  | Both the context of the evaluation and the intervention logic that was developed by the evaluation team are clearly and thoroughly described. Explanation of general constraints was focused on the fact that the evaluation was conducted in parallel with work on development of the next Strategic Plan and what was done about that. |      |
| 3. Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change, and assess the adequacy of these?   | Yes                  | Together the description in the report and the Evaluation Matrix in Annex 3 clearly present the evaluation approach and framework as well as a description of methods chosen. Figure 1 in the report and Annex 2   |      |
| <i>To ensure a rigorous design and methodology</i>   |                      | present the intervention logic. It is evident from the report, that a reconstructed logic was used to derive the evaluation questions and design the evaluation matrix.  |      |
| 4. Is the evaluation framework clearly described in the text and in the evaluation matrix? Does the evaluation matrix establish the evaluation questions, assumptions, indicators, data sources and methods for data collection? | Yes                  | The evaluation matrix is concise, being 6 pages. It covers the required items: evaluation questions, indicators, sources of information, methods, and tools for the data collection. Several shortfalls are evident: the methods for analysis are not clearly described for all types of data although                                   |      |
| 5. Are the tools for data collection described and their choice justified?   | Yes                  | Chapter 1.3. explains data collection techniques. Limitations and mitigating measures are described very briefly in the section 1.4. For instance, it is said in the annex 8 that another limitation was time constraints (p. 48).   |      |
| 6. Is there a comprehensive stakeholder map? Is the stakeholder consultation process clearly described (in particular, does it include the consultation of key stakeholders on draft recommendations)?                           | Yes                  | The sampling strategy is described in the annexes 8, 11 and 12. The intent to involve stakeholders as set forth in the ToR was confirmed explicitly with a thorough description of stakeholder involvement early in the report proper (Chapter 1).   |      |
| 7. Are the methods for analysis clearly described for all types of data?   | Yes                  | The methodology includes the declaration statement that the "evaluation met the UNEG gender-related norms and standards to the extent possible and as appropriate for this type of evaluation" (p. 2), but the evaluators  |      |
| 8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)   | Yes                  | do not explain how they implemented those norms and standards. They did note that assessing cross-cutting issues on SWAP/ GEEW did not appear appropriate or necessary given the nature of this evaluation.  |      |
| 9. Is the sampling strategy described?   | Yes                  |  |      |
| 10. Does the methodology enable the collection and analysis of disaggregated data?   | Partial              |  |      |
| 11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?   | Partial              |  |      |

|   |                      |  |                  |
|---|----------------------|--|------------------|
| <b>3. Reliability of Data</b>   | Yes<br>No<br>Partial | Assessment Level:  | <b>Good</b>      |
| <i>To ensure quality of data and robust data collection processes</i>   |                      | Comment: Available data sources were effectively utilized and the quality of data is ensured through triangulation. The report notes that collected data, both qualitative and quantitative, was triangulated "by source of evidence/data collected" (p. 4). There is clear evidence of the triangulation.   |                  |
| <b>1. Did the evaluation triangulate data collected as appropriate?</b>   | Yes                  | The report noted possible limitations in primary and secondary sources but did not explain what was done to minimize such issues. Constraints, limitations, and mitigating measures are described very briefly in the section 1.4. One issue not addressed was that the rate of response to the surveys was relatively low, ranging from 62 percent for country offices to 20 percent for national counterparts and member states. The distribution of respondents is well-described, but the extent to which this is representative of the universe is not indicated. |                  |
| <b>2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?</b>   | Yes                  | Annex 11 shows that surveys were confidential, but there are no explanations in the methodology and annexes how the issues of discrimination and other ethical considerations were considered in conducting country and regional studies and interviews with key informants (pp. 4-5).   |                  |
| 3. Did the evaluation make explicit any possible limitations (bias, data gaps etc.) in primary and secondary data sources and if relevant, explained what was done to minimize such issues? | Yes                  |  |                  |
| 4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?  | Partial              |  |                  |
| <b>4. Analysis and Findings</b>   | Yes<br>No<br>Partial | Assessment Level:  | <b>Very good</b> |
| <i>To ensure sound analysis and credible findings</i>   |                      | Comment: 1. The data analysis reflects contextual factors and the findings are concise and well substantiated by evidence. The organization of this section of the report is greatly facilitated by the fact that it is structured according to the five evaluation questions with core findings supported by evidence collected across sources.   |                  |
| <b>1. Are the findings substantiated by evidence?</b>   | Yes                  | There are references to the interviews, documents in the text and supportive annexes (documentation, country and regional studies, administrative data analysis, and survey results). This greatly helps the presentation of the analysis and interpretation in a systematic, logical and consistent way.  |                  |
| 2. Is the basis for interpretations carefully described?  | Yes                  | 2. Interpretations are based on carefully described assumptions - an example is the statement that "counting outputs... is a useful proxy for focus but has its limitation in hiding fragmented activities... These, by themselves, may not necessarily be bad and may reflect an appropriate response to the national context" (p. 21).   |                  |
| 3. Is the analysis presented against the evaluation questions?  | Yes                  | 3. As noted, the analysis in the Findings section is presented against the evaluation questions quite effectively.   |                  |
| <b>5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?</b>   | Yes                  | 4. The analysis is transparent about sources and quality of data - often citing for example "country counterparts felt" and "the country studies suggest..." and "differing needs of country offices" (p. 31).   |                  |
| 6. Does the analysis show different outcomes for different target groups, as relevant?  | Yes                  |  |                  |
| <b>7. Is the analysis presented against contextual factors?</b>   | Yes                  |  |                  |

|   |            |   |
|---|------------|---|
| <p>8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?</p> | <p>Yes</p> | <p>5. The cause-effect links are explained in the text as seen in the treatment of how the introduction of differentiated modes of engagement were introduced without adequate guidance and therefore did not support the staff behavioral changes needed. Another example was the statement that "the focus on policy/advocacy can be a hindrance to resource mobilization" (p. 39). In the meantime, cause and effect links are not explained in the Annex 2 which contains theories of change. The diagrams have clear structure, but the general diagram has too many confusing links between 3 groups of strategic elements, 14 outputs, and 5 outcomes, because outputs are not combined into small groups.. Unintended outcomes are not highlighted in the way to easily to find them.</p> <p>6. Analysis is disaggregated to show different outcomes even within target groups, whereby countries interviewed differed by color of quadrant from one another in the priority they gave to it on the need to enhance policy work versus responding to new or emerging needs (p. 30). The analysis shows different results for different countries and implementing partners (annex (c) "Survey results", and annex 11 "Administrative data analysis").</p> <p>7. The report does a very good job of presenting analysis against contextual factors: global, humanitarian, UN development system contexts (p. 13).8. The report states inequality (gender) is factored in (p 19) and the focus on gener and rights are maintained and strengthened through the architecture (p 44) The analysis elaborates on cross-cutting issues in the following sections: 1.2 "Overall evaluation framework" (p. 2), 2.1 "Background to the design of the strategic plan 2014-2017" (pp. 7-8), chapter 3 "Findings" (p. 18, p. 32, p. 40, p. 44).</p> |
|---|------------|---|

|  |     |  |                   |                  |
|--|-----|--|-------------------|------------------|
| <b>5. Conclusions</b>  |     | Yes<br>No<br>Partial   | Assessment Level: | <b>Very good</b> |
| <i>To assess the validity of conclusions</i>   |     | Comment: Conclusions are linked to findings and are organized by evaluation questions.   |                   |                  |
| <b>1. Do the conclusions flow clearly from the findings?</b>   | Yes | 1. The conclusions are based on and logically flow from the findings. The conclusions are cross-referenced to the specific findings in the previous section (e.g., cf p. 45)<br>2. The conclusions focus at the macro level on three strategic key messages that UNFPA management needs to address in improving the support architecture.<br>3. Conclusions seem to convey the evaluators' unbiased judgment.  |                   |                  |
| 2. Do the conclusions go beyond the findings and provide a thorough understanding of the underlying issues of the programme/initiative/system being evaluated?                 | Yes |  |                   |                  |
| 3. Do the conclusions appear to convey the evaluators' unbiased judgement?   | Yes |  |                   |                  |
| <b>6. Recommendations</b>  |     | Yes<br>No<br>Partial   | Assessment Level: | <b>Very good</b> |
| <i>To ensure the usefulness and clarity of recommendations</i>   |     | Comment: 1. Recommendations flow logically from conclusions and are in fact cross-referenced at the beginning of this section.   |                   |                  |
| <b>1. Do recommendations flow logically from conclusions?</b>  | Yes | 2. The recommendations are clear and targeted to the appropriate unit of UNFA; whether they are operationally feasible is not immediately self-evident given they are formulated in a manner best understood by internal staff, although the management response suggests that they are feasible.<br>3. The recommendations reflect stakeholder consultations that culminated in a validation workshop in February 2017.<br>4. While there are only a total of 7 recommendations, there also are "specific operational suggestions" which provide another avenue for further alternative actions. The recommendations suggest a timetable for implementation in the Strategic Plan 2018-2021.<br>5. The recommendations are prioritized and clearly presented. |                   |                  |
| <b>2. Are the recommendations clearly written, targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?</b> | Yes |  |                   |                  |
| 3. Do recommendations appear balanced and impartial?   | Yes |  |                   |                  |
| 4. Is a timeframe for implementation proposed?   | Yes |  |                   |                  |
| 5. Are the recommendations prioritised and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?                      | Yes |  |                   |                  |

| 7. Gender  | 0<br>1<br>2<br>3 (**) | Assessment Level:  | Fair |
|--|-----------------------|--|------|
| <i>To assess the integration of Gender Equality and Empowerment of Women (GEEW) (*)</i>  |                       | Comment: While gender issues were discussed under findings, conclusions and recommendations do not reflect a gender analysis. The evaluators do not explain how they applied UNEG guidance on gender equality and human rights, even though they indicate that they followed them.   |      |
| 1. Is GEEW integrated in the evaluation scope of analysis and indicators designed in a way that ensures GEEW-related data to be collected?   | 2                     | 1. Scope and indicators.   |      |
| 2. Do evaluation criteria and evaluation questions specifically address how GEEW has been integrated into design, planning, implementation of the intervention and the results achieved? | 1                     | 1.2. The evaluation matrix includes GEEW indicators, for instance, "SRH and GBV are adequately taken into account in national preparedness and contingency plans in high risk countries" (p. 20).<br>1.3. The evaluators did not plan to collect the evidences on how GEEW results were mainstreamed into the program themes.  |      |
| 3. Have gender-responsive evaluation methodology, methods and tools, and data analysis techniques been selected?   | 1                     | 2. Evaluation criteria and evaluation questions  |      |
| 4. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?  | 2                     | The evaluation criteria and evaluation questions do not address how GEEW has been integrated into design, planning, implementation of the intervention and the results achieved.<br>3. Methodology.<br>3.1. Though the data itself was triangulated, it's not clear whether the voices of both men and women are reflected.<br>3.2. Data was partially disaggregated by sex in the following annexes: List of persons interviewed (f) and (b) List of stakeholders interviewed.<br>3.2. Gender-responsive approach is considered, but it is unclear how it was incorporated into the design and implementation of the evaluation.<br>3.3. The evaluators do not explain how they applied UNEG guidance on gender equality and human rights, even though they claim that they followed them (p. 2).<br>4. Findings, conclusions, and recommendations<br>4.1. We can see from the report that the evaluators tried to assess progress on GEEW wherever possible: "The 2016 World Humanitarian Summit (WHS) produced a number of commitments to improve the response in humanitarian situations (including empowering and protecting women and girls)" (p. 13). They analyzed signature indicators, including "women and/or girls" (p. 40).<br>4.2. In general, the evaluation conclusions reflect a gender analysis: "The focus on gender and rights were also maintained and strengthened through the architecture (for example in the RAS and the integrated results framework)" (p. 44). There are no specific GEEW conclusions.<br>4.3. Evaluation recommendations do not reflect a gender analysis. |      |

(\*) This assessment criteria is fully based on the UN-SWAP Scoring Tool. 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).

(\*\*) Scoring uses a four point scale (0-3).

0 = Not at all integrated. Applies when none of the elements under a criterion are met.

1 = Partially integrated. Applies when some minimal elements are met but further progress is needed and remedial action to meet the standard is required.

2 = Satisfactorily integrated. Applies when a satisfactory level has been reached and many of the elements are met but still improvement could be done.

3 = Fully integrated. Applies when all of the elements under a criterion are met, used and fully integrated in the evaluation and no remedial action is required.

| 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                        |  |
| <b>Total scoring points</b>  | <b>69</b>                             | <b>24</b>                | <b>7</b>                 |  |
| <b>Overall assessment level of evaluation report</b>                   | <b>Very Good</b>                      |                          |                          |  |
|  | Very good<br>very confident<br>to use | Good<br>confident to use | Fair<br>use with caution | Unsatisfactory<br>not confident to use |

- (\*) (a) Insert scoring points associated with criteria in corresponding column (e.g. - if 'Analysis and findings' 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 explain

- How it can be used?

- What aspects to be cautious about?

**Where relevant, please explain the overall assessment Very good, Good or Unsatisfactory**

This is a very thorough evaluation of a complex subject since results flowing from the architecture of the organization are not easy to assess. However, the evaluation did a thorough job of examining progress and factors that could be improved in future planning.

**Consideration of significant constraints**

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

Yes

No

If yes, please explain: