

Organizational unit: UNFPA **Year of report:** 2017

Title of evaluation report: Country Programme Evaluation: India. Eighth Programme Cycle (2013 - 17) Evaluation report

Overall quality of report: **Very Good** **Date of assessment:** 19 September 2017

Overall comments: This is a very thorough and clearly written country programme evaluation. The evaluators presented a reconstructed logic for the overall program and each intervention, as well as an explanation for each of the suggested changes. In addition to document review, interviews, and group discussions, the CPE used a case study approach to identify the main characteristics of initiatives that had demonstrated high impact. This was particularly useful for identifying lessons learned. A solid overview of the findings was presented in the main text and meticulously documented in an extensive evaluation matrix in the annex. The conclusions flowed from the findings and recommendations flowed from the conclusions - and both were clear. The evaluation had a strong gender focus. There was a good examination of how GEEW, including gender mainstreaming, was addressed in each programme area, and there are well justified conclusions and recommendations specific to GEEW. Shortcomings included the lack of gender disaggregated data.

Assessment Levels

- Very Good** (blue box): strong, above average, best practice
- Good** (green box): satisfactory, respectable
- Fair** (yellow box): with some weaknesses, still acceptable
- Unsatisfactory** (red box): 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> | | |
|---|---|--|------------------|
| 1. Structure and Clarity of Reporting | Yes No Partial | Assessment Level: | Very good |
| <i>To ensure the report is comprehensive and user-friendly</i> | | | |
| 1. 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? | Yes | This report has a clear structure. It is logically organized and easy to read. | |
| 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 | The text of the main report just meets the max length for CPEs, excluding the Executive Summary it is 70 pages. The annex is a separate document and is 159 pages, mainly due to the inclusion of very detailed findings notes within the Evaluation Matrix. | |
| 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)? | Yes | The report is structured logically. It does not include a specific section on lessons learned, however lessons are integrated into the Findings, and in particular the case study results in section 4.6 (p 42). | |
| 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 | The annexes are complete. The Annex 2 includes description of the ERG members and minutes from the meetings which could be considered as evidence of stakeholder consultation process. | |
| <i>Executive summary</i> | | | |
| 5. Is an executive summary included in the report, written as a stand-alone section and presenting the main results of the evaluation? | Yes | The Executive Summary is written as a stand-alone document and it describes the main results of the evaluation. | |
| 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 | The Executive Summary follows the required structure, but while the purpose of the evaluation is not clearly mentioned the text stated that the evaluation was conducted “to inform the development of the next programming cycle” (p x) | |

| | | |
|---|-----|---|
| 7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)? | Yes | The Summary is concise and is just under 5 pages in length. |
|---|-----|---|

| | | | |
|----------------------------------|----------------------|-------------------|-------------|
| 2. Design and Methodology | Yes No Partial | Assessment Level: | Fair |
|----------------------------------|----------------------|-------------------|-------------|

To ensure that the evaluation is put within its context

| | | |
|---|----|---|
| 1. Does the evaluation describe the target audience for the evaluation? | No | The methodology does not describe the target audience for the evaluation. |
|---|----|---|

| | | |
|---|-----|--|
| 2. Is the development and institutional context of the evaluation clearly described and constraints explained? | Yes | Country context is well explained in the chapter 2: Economic and Social Context, Population dynamics, Reproductive health and rights, Gender equality, and Disaster. |
|---|-----|--|

| | | |
|---|-----|---|
| 3. Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change, and assess the adequacy of these? | Yes | The final report describes the reconstruction of the logic in the section 3.4 "Reconstructing the CPAP-8 Intervention Logic." The consultants assessed original intervention logic under each thematic area, explained the rationale for the reconstruction of the intervention logic, and provided the reconstructed intervention logic, presented in diagram form on page 23. Annex 8 includes a detailed evaluability assessment of each output indicator. |
|---|-----|---|

To ensure a rigorous design and methodology

| | | |
|---|-----|--|
| 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 framework is designed in accordance with the UNFPA requirements (p.1). The evaluation questions are clear and correspond to the OECD–DAC evaluation criteria, including UNFPA criteria "Added Value". Figure 1.3 maps evaluation questions according to the evaluation criteria. The evaluation matrix is found in the annex 7, it is designed in accordance with the UNFPA Handbook for evaluation at UNFPA (pp. 109-110) and meets the sub-criterion. |
|---|-----|--|

| | | |
|--|---------|---|
| 5. Are the tools for data collection described and their choice justified? | Partial | The tools for data collection are clearly explained: documents review, semi-structured personal interviews, group interviews, and field visits (pp. 4-5). The tools are described but not explicitly justified. A key source of information was four recently conducted thematic evaluations. |
|--|---------|---|

| | | |
|--|-----|--|
| 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 | There is no comprehensive stakeholder map, but the Table 1.1. describes categories of stakeholders for individual and group interviews (p. 6). Even though there is no separate paragraph devoted to explanation of the stakeholder consultation process, it is possible to understand it from the text. For instance, the evaluators explain that "At all stages of the evaluation, semi-structured group interviews were conducted with external stakeholders, including counsellors, consultants, and government or NGO programme staff at state level to respond to a variety of evaluation questions" (p. 5). The Evaluation Reference Group was established to validate the findings, including the recommendations "The purpose of the ERG meeting was to present preliminary findings and of conclusions and recommendations of the UNFPA country programme evaluation" (Annex 2). |
|--|-----|--|

| | | |
|--|---------|---|
| 7. Are the methods for analysis clearly described for all types of data? | Partial | The methods for analysis included "...the exercise of matching relevant portions of the interview transcripts with the corresponding evaluation question on the matrix and consulting some additional documentary sources in order to write the final report" (p. 5). This is a very general description of the analysis. |
|--|---------|---|

| | | |
|--|---------|--|
| 8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?) | Partial | The limitations presented included a short time frame for the complexity of the evaluation and several challenges to accessing data sources. Actions to overcome each are provided. However bias is not addressed and should have been in respect to the representativeness and illustrativeness of the stakeholders consulted. |
| 9. Is the sampling strategy described? | Yes | The evaluators preferred purposive sampling of interviewees (p. 3). The sampling strategy is described in the section "Selection of the sample of stakeholders" (pp. 5-6). Table 1.1 presents the number of interviews that were undertaken during the evaluation process. |
| 10. Does the methodology enable the collection and analysis of disaggregated data? | Yes | The methodology (evaluation questions, thematic assessments, interview questions) enables the collection and analysis of disaggregated data. |
| 11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)? | Yes | The methodology is appropriate for assessing cross-cutting issues. Examples demonstrating this include that the consultants explain that "Cross-cutting issues of vulnerability, gender mainstreaming, resource mobilization, corporate social responsibility and south-south collaborations were also addressed in the data collection" (p. 3); "A small number of high-level stakeholders were chosen to address cross-cutting issues of resource mobilization, corporate social responsibility and south-south collaboration" (p. 6). |

| | | | |
|-------------------------------|----------------------|-------------------|-------------|
| 3. Reliability of Data | Yes No Partial | Assessment Level: | Good |
|-------------------------------|----------------------|-------------------|-------------|

To ensure quality of data and robust data collection processes

| | | |
|--|-----|---|
| 1. Did the evaluation triangulate data collected as appropriate? | Yes | The evaluation consultants triangulated data appropriately: "Methods to triangulate data included review of documents from different sources and purposive stakeholder sampling to represent varying perspectives" (p. x). It is possible to find examples of triangulation in the text: "Analysis of the census data and publications on the skewed sex ratio at birth ..." (p. 36), "Programmes or initiatives were found to make the best use of UNFPA resources ... Stakeholders also pointed that the continuity of UNFPA's commitments to topics and to partnerships is a source of value added" (p. 47). |
| 2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources? | Yes | The data sources (documentation; semi-structured individual and group interviews with key informants; four in-depth thematic assessments carried out the month before the CPE; field visits to programmes in the five states; interviews with UNFPA country and state staff, and a case-study method) are justified by the Annex 2 "ERG Members & Minutes," the Annex 3 "List of documents consulted," the Annex 4 "Persons Interviewed," and the Annex 9 "Interviews Guide and Template." This indicates that the data are reliable. |
| 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 | The limitations, primarily gaps in available program data, were clearly explained. For instance, the Ministry of Health and Family Welfare did not participate in the initial ERG, the consultants faced a lack of state-level strategic plans, necessary activities were not entered into the UNFPA GPS system, there were issues with the use of the publications produced in CP-8 (p. 7). The team's approach to addressing gaps included seeking supplementary information from staff and further document review. |

| | | |
|--|---------|--|
| 4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations? | Partial | While the report does not explicitly explain how UNFPA and UNEG ethical standards were used, there is a note that “The final assessment reports were made available to the evaluation team, and a bilateral sharing of interview notes and data was also facilitated without the involvement of UNFPA to maintain strict ethical considerations of the evaluation process” (p. 4). This suggests that the expected sensitivity was maintained. |
|--|---------|--|

| | | | |
|---------------------------------|----------------------|-------------------|------------------|
| 4. Analysis and Findings | Yes No Partial | Assessment Level: | Very good |
|---------------------------------|----------------------|-------------------|------------------|

To ensure sound analysis and credible findings

| | | |
|---|-----|--|
| 1. Are the findings substantiated by evidence? | Yes | The evidence for the findings are indicated in footnotes and explained more fully in the Evaluation Matrix in the annex. The evaluators refer to the evidence while explaining their findings. For instance, they state that findings on Major A/Y Programmes are based on such sources of data as “CPE and thematic assessment interviews with RKSK consultants and lead government officials in Delhi, Bihar, Rajashtan, and MP, where Jejeebhoy & CPE team member visited an AFHC clinic. There is no information from Odisha, and Maharashtra RKSK consultants are supported by UNICEF” (p. 24). |
|---|-----|--|

| | | |
|--|-----|---|
| 2. Is the basis for interpretations carefully described? | Yes | <p>The basis for interpretations is carefully described throughout the analysis by comparison of different sources of data (references to the interviews and documents) and description of specific cases. In the following example, the evaluators are very accurate in their interpretations “Several CP-8 RH/FP interventions have a high potential for impact, but results could not be verified in field visits” (p. 29). In other cases the consultants have demonstrated that “Research on assisted reproductive technology (ART) and commercial surrogacy as well as data collection on hysterectomy and female genital mutilation provides evidence for policy and programme changes for protecting reproductive rights ...(p. 18); “the GBSS case study showed a progression over several country programmes...” (p. 39).</p> <p>In the meantime, there are places in the report where more specific data could be provided. For instance, the following finding could have been strengthened by mentioning a couple of examples from the end-line assessment and explaining the reasons of no follow-up support: “The end-line assessment showed improvements in capacity and capability; however, no follow-up support has been provided since 2015” (p. 31).</p> |
|--|-----|---|

| | | |
|--|-----|---|
| 3. Is the analysis presented against the evaluation questions? | Yes | The section is well structured with the evaluation question presented before a concisely written summary of findings and then the more detailed findings. |
|--|-----|---|

| | | |
|---|-----|---|
| 4. Is the analysis transparent about the sources and quality of data? | Yes | The analysis is transparent about the sources and quality of data. The sources of data are usually mentioned while describing the findings like interviews with state FP officers (p. 29), IIPS & UNFPA (2015) Conditional Cash Transfers for Girls in India (p. 34), Jejeebhoy, S, Kulkarni, et al., Health and Population in India (p. 39). It is also possible to find examples of comments on the quality of data, for instance, the consultants note that “Saathiya share technically accurate information on adolescent issues with their peers” (p. 25). |
|---|-----|---|

| | | |
|--|-----|--|
| 5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted? | Yes | This evaluation is careful to show cause and effect links between the output of the programmes and projects and what was expected to happen. In addition to highlighting the successes, the shortfalls, mixed results and missed opportunities are also clearly explained. In a few cases there is a lack of details. For instance, the following example demonstrates that “policy guidance to ... safe and legal abortions” (output) was “used by MOHFW and in states implementing the Act” (outcome) – page 29. In another example, the evaluators reveal that study “Assessment of Tribal Sub-Plan and Scheduled Caste Sub-Plan Programme Implementation In Odisha” (output) has contributed “directly into building institutional capacity for social inclusion through use of the study and its tools to monitor progress in several departments... (p. 36). |
| 6. Does the analysis show different outcomes for different target groups, as relevant? | Yes | The results for different target groups are shown, for example the A/Y programmes are noted as being successful in many cases but not in reaching married girls and early drop-outs. The evaluators also look at the results on a state-by-state basis which is important for as large and diverse a country as India. |
| 7. Is the analysis presented against contextual factors? | Yes | The analysis is presented against contextual factors which are well described in the report, for instance: “the political, economic, and socio-cultural context of each state is unique, and in the federal system, state governments enjoy considerable autonomy” (p. 41). |
| 8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights? | Yes | The analysis includes cross-cutting issues. For example, the discussion on FP (p 30) notes that consultants have not been particularly effective in advocating for rights-based approaches, or for greater gender mainstreaming. There is also a specific evaluation question addressing gender bias sex selection and gender mainstreaming. The consultants state that they addressed cross-cutting issues of vulnerability, gender mainstreaming, resource mobilization, corporate social responsibility and south–south collaborations through “complementary review of documents, purposive sampling of interviewees and interactions with development partners and UN agencies” (p. 3). |

| | | | |
|--|----------------------|--|------------------|
| 5. Conclusions | Yes No Partial | Assessment Level: | Very good |
| <i>To assess the validity of conclusions</i> | | | |
| 1. Do the conclusions flow clearly from the findings? | Yes | Conclusions flow clearly from the findings. All the conclusions are linked to the evaluation questions, for instance, conclusion I relates to the EQ#5 and #8 (p. 56). Also, the consultants support their statements with evidence from the findings. For instance, the report states that “Effective strategies and technical expertise developed in one state have maximum impact when also used for capacity-building in other states and at national level” (p. 62). Also, the evaluators discuss the following example: “technical expertise in GBSS in Maharashtra -- acquired over several years of programming -- was employed effectively to build capacity in other UNFPA states and beyond” (p. 62). | |
| 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 | The consultants provide a thorough explanation of the underlying issues of the programme: gaps, cause and effect links, recourse value, and others. For instance, the conclusion 3 relates to “UNFPA’s ability to mobilize high quality technical assistance (TA) and expertise for national priorities” (p. 60). The consultants state that “UNFPA’s comparative advantage in all thematic areas depends on its ability to respond to national and state needs” like when “TA consultants with proven high-level expertise were chosen” (p. 60). | |

| | | |
|--|-----|--|
| 3. Do the conclusions appear to convey the evaluators' unbiased judgement? | Yes | Conclusions appear to convey the evaluators' unbiased judgement: there are examples from the findings, references to the evaluation questions and other conclusions. But, in some cases the conclusions lack of direct references to the findings like "CPE evidence suggests" (too general) that "when pilot community-based programmes rely on additional personnel rather than building capacity in existing human resources in government programmes, the pilots are less likely to be scaled-up in their current form" (p. 64). |
|--|-----|--|

| | | | |
|---------------------------|----------------------|-------------------|-------------|
| 6. Recommendations | Yes No Partial | Assessment Level: | Fair |
|---------------------------|----------------------|-------------------|-------------|

To ensure the usefulness and clarity of recommendations

| | | |
|---|---------|---|
| 1. Do recommendations flow logically from conclusions? | Yes | Recommendations follow conclusions: one after another and are based on the conclusions. |
| 2. Are the recommendations clearly written, targeted at the intended users and action-oriented (with information on their human, financial and technical implications)? | Partial | They are clearly written and action oriented. They do not highlight the intended user, although it can be assumed that the main audience is senior country-level staff. In most cases the operational implications are identified, but not financial implications. For instance, recommendation 4 specifies that "a full-time M&E staff person with expertise in both quantitative and qualitative methods and experience in process evaluation." Technical and financial implications include "...developing a network of qualified evaluation consultants to meet specific needs, and including M&E costs in implementing partner budgets as needed" (p. 62). |
| 3. Do recommendations appear balanced and impartial? | Yes | The recommendations appear balanced and impartial. |
| 4. Is a timeframe for implementation proposed? | Yes | The introduction to the recommendations suggests that all the recommendations should be considered for the development of the CP-9. |
| 5. Are the recommendations prioritised and clearly presented to facilitate appropriate management response and follow up on each specific recommendation? | Yes | Recommendations are clearly presented to facilitate appropriate management response and follow up on each specific recommendation. The recommendations are well structured and logical. The recommendations are prioritized (Medium to High), but most recommendations have High priority. |

| | | | |
|------------------|------------------|-------------------|-------------|
| 7. Gender | 0 1 2 3 | Assessment Level: | Good |
|------------------|------------------|-------------------|-------------|

To assess the integration of Gender Equality and Empowerment of Women (GEEW) ()*

| | | |
|--|---|--|
| 1. Is GEEW integrated in the evaluation scope of analysis and indicators designed in a way that ensures GEEW-related data to be collected? | 3 | GEEW was included into the evaluation scope. The evaluators assessed four thematic areas: reproductive health and family planning, adolescents and youth, gender, and population dynamics (p. 1). The evaluation matrix has GEEW indicators, for instance: "gender-sensitive knowledge" (p. 66), "incorporating gender sensitivity and youth-friendly approaches" (p. 71). |
| 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? | 2 | Evaluation criteria are based on OECD-DAC criteria: relevance, effectiveness, efficiency, and sustainability. The evaluation focused on added value as an additional criteria (p. 3). The evaluation criteria do not address GEEW specifically. EQ3 has a gender component: "In what ways has UNFPA supported gender equality and to what extent has it contributed to: (i) improved responses to gender biased sex selection and (ii) gender mainstreaming across the programming area? (p. 3). |

| | | |
|--|---|--|
| 3. Have gender-responsive evaluation methodology, methods and tools, and data analysis techniques been selected? | 2 | <p>Methodology and interview guides do not explain how the evaluation consultants implemented ethical standards.</p> <p>The methodology does not provide specific approaches for fostering participation and inclusiveness (recommended by the table 6.7. in the UNEG Guidance Integrating Human Rights and Gender Equality in Evaluations). But, the report presents gender-responsive data when necessary, for instance, the consultants found out that “merely 31 percent men and 15 percent women above the age of 20 years in India have completed high school education and gender disparities persist even amongst the young” (p. 8).</p> <p>A desk review and data analysis stages included gender analysis which is evident from the annex 3 “List of Documents Consulted” and Annex 4 “Persons Interviewed.”</p> |
| 4. Do the evaluation findings, conclusions and recommendations reflect a gender analysis? | 3 | <p>Evaluation findings reflect a gender analysis. It is possible to find gender-responsive data in the different sections of the report: Effectiveness in the Adolescence and Youth Programmatic Area (for instance, gender-sensitive life skills – page 23), Effectiveness in the Reproductive Health and Family Planning (for instance, gender-based violence - page 31), Effectiveness in the Gender Biased Sex Selection and Gender Mainstreaming (for instance, Gender Biased Sex Selection - page 32), Effectiveness in Population Dynamics Programmatic Area (for instance, gender and social inclusion – page 35). Evaluation conclusions reflect a gender analysis, for instance: conclusion 2 relates to the gender biased sex selection (p. 57), conclusion 5 states that “Cross-state learning would be particularly valuable for the large multi-state programmes in A/Y and gender” (p. 62), the conclusion 6 considers gender issues in explaining “design and implementation of large national and state adolescent health and life-skills programmes” (p. 64), conclusion 8 addresses gender inequality (p. 67), and conclusion 9 focuses on gender issues as well (p. 68). Evaluation recommendations reflect a gender analysis, for instance, recommendation 2 describes gender mainstreaming issues (p. 59), recommendation 6 refers to GBV (p. 65), and recommendation 9 addresses gender discrimination (p. 68).</p> |

(*) 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).

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 | 18 | 24 | |
| 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 '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?

The fair rating is based on unevenness in the evaluation design and in its applications in terms of findings. While the weaknesses of the methods and findings need to be considered, the fact that the conclusions are strong and clearly expressed, suggests that the evaluation can be used to help design the next country program.

- What aspects to be cautious about?

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

Consideration of significant constraints

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

Yes No

If yes, please explain: