



Organizational unit: **India CO**

Year of report: **2022**

Title of evaluation report: **EVALUATION OF THE UNFPA NINTH COUNTRY PROGRAMME OF ASSISTANCE TO THE GOVERNMENT OF INDIA CP9 (2018-2022)**

Overall quality of report: **Very good**

Date of assessment: **1 June 2022**

**Overall comments:** This is a sound final evaluation of the ninth country programme in India with the purpose of determining what should be incorporated in the next country programme. It was complex to conduct given the size of the country programme and the restrictions of the covid pandemic. The evaluation was based on a thorough review of many documents (124) and interviews and focus groups with 899 persons. The methodology is well described with the approach including field visits to a systematically selected set of locations in the four provinces in which UNFPA was working. Using a carefully developed (and revised) theory of change, as well as complete evaluation questions structured around UNEG criteria, the evaluation collected extensive data to determine results of the UNFPA interventions and what worked and what did not. Included was a careful effort to address gender issues. The findings were extensive (36) and well documented and triangulated, although page limitations meant that many were reflected in detailed annexes. Unanticipated outcomes were included in the analysis. Although the recommendations could have been more clearly presented, they were sound and practical. Given the complexity of the CP and the extensive amount of data collected, the evaluators were challenged to keep the document within the standard page limit. One result was the extensive use of abbreviations (over 125) to reduce length but which made reading difficult. Readability could also be improved by more use of visual aids and design elements to highlight main points, and by combining the annexes into one document for easier reference. The evaluation was disability inclusive. Evaluators were purposeful in looking at results relating to disability even though these were limited because this element was added after the CP had been defined and there were issues of access to PWD because of COVID-19. Disability-related questions were included in interview protocols where relevant. Disability issues were also covered in conclusions and recommendations.

**Assessment Levels**

<b>Very Good</b> (blue box)	strong, above average, best practice	<b>Good</b> (green box)	satisfactory, respectable	<b>Fair</b> (yellow box)	with some weaknesses, still acceptable	<b>Unsatisfactory</b> (red box)	weak, does not meet minimal quality standards
-----------------------------	--------------------------------------	-------------------------	---------------------------	--------------------------	--	---------------------------------	---

Quality Assessment Criteria	Insert <i>assessment level</i> followed by main <i>comments</i> . (use 'shading' function to give cells corresponding colour)	
<b>1. Structure and Clarity of Reporting</b>	Yes No Partial	Assessment Level: <b>Fair</b>
<i>To ensure the report is comprehensive and user-friendly</i>		
<b>1. Is the report structured in a logical way? 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? Is there a clear distinction made between analysis/findings, conclusions, recommendations and lessons learned (where applicable)?</b>	Partial	The report follows the standard structure, uses rights-based terminology where relevant, and includes a section on lessons learned. This is a very complex evaluation and a defect is that there is an immense use of abbreviations and acronyms that make the text sometimes difficult to read. The table on abbreviations shows 114 abbreviations and acronyms, but these do not include all that are used. For example, IP is not included and it isn't until Finding 30 that IP is shown as "implementing partner". There is adequate use of visual aids in the preliminary sections but the findings are quite text heavy and are only broken up by a few photos (these are a good addition but are quite small). More attention to formatting including use of larger font for section headings, pull-out boxes for participants quotes, more bolded text to highlight the topic of paragraphs (including for the listing of lessons learned), more consistent spacing (see, for example, the Conclusions section), and additional graphics would increase readability. There are also a number of spelling errors, although these do not affect the readability of the report.
<b>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)</b>	Partial	The main report including the Executive Summary is 80 pages. Since this is a very complex country programme, a Partial is given.
<b>3. Do the annexes contain – at a minimum – the ToRs; a bibliography; a list of interviewees; the evaluation matrix; methodological and data collection tools used (e.g. interview guides; focus group notes, outline of surveys)?</b>	Partial	The annexes contain the required elements plus additional annexes on performance data, findings for each programme area, the sampling process, two case studies on gender, and an overview of the State Programme. The annexed material is appropriate but could have been more professionally presented - each annex is in a separate file (some in pdf and some in word format), and some were in draft form. Good practice is for annexes to be well formatted and combined into a single document (which can be a separate volume from the main report).
<i>Executive summary</i>		
<b>4. Is an executive summary written as a stand-alone section, presenting the i) Purpose; ii) Objectives, scope and brief description of interventions; iii) intended audience; iv) Methodology; v) Main results; vi) Conclusions and Recommendations?</b>	Yes	The summary contains the expected material from the sections and combines conclusions and recommendations.
<b>5. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?</b>	Partial	The executive summary is 4.5 pages which is within the expected length. However, the presentation would be clearer if more subheadings or bullet points were used, particularly for the findings/conclusions and recommendations sections.
<b>2. Design and Methodology</b>	Yes No Partial	Assessment Level: <b>Very good</b>

<i>To ensure that the evaluation is put within its context</i>		
1. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	The context is well-described, including the constraints occasioned by COVID-19.
2. Does the evaluation report discuss and assess the intervention logic and/or theory of change?	Yes	The evaluators made a thorough effort to improve on the theory of change and used the reconstructed theory, which was discussed with the Country Office, in the collection of data.
<i>To ensure a rigorous design and methodology</i>		
3. 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	A very detailed evaluation matrix was developed, including all of the evaluation questions organized by what are called "outputs" (which are really objectives based on type of work) and within them by criteria with efficiency and sustainability common to all three outputs. It shows the assumptions, indicators, data sources and methods for data collection, but does not include summarized findings. The latter can be found in other annexes that address performance for each main programme area.
4. Are the tools for data collection described and their choice justified?	Yes	There is a complete description of data collection tools: a standard combination of secondary sources based on documents and their analysis and primary sources consisting of interviews with stakeholders, who were described in a stakeholder map, focus groups and observation in a selection of sites. In each case, the basis for selection were described in detail and reflected also in annexes.
5. 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	A stakeholder map was created and is reflected in detail in an annex. The stakeholder process is carefully described, noting particularly that the first draft report was shared for review by CO staff, ERG, APRO and Evaluation Manager for feedback. The final draft, updated upon taking the feedback into consideration, was shared with the national stakeholders and CO staff for validation. (p. 10)
6. Are the methods for analysis clearly described for all types of data?	Yes	There is a section on data analysis that described how each of the data sources was added into the analysis. The main method was content analysis where triangulation was used to "interpret qualitative data with some quantitative analysis using secondary data."
7. Are methodological limitations acknowledged and their effect on the evaluation described? Does the report discuss what was done to minimize such issues?	Yes	There is a thorough discussion of limitations, many of which were caused by COVID-19, which made interviews with students, for example, difficult. There was a discussion of methods to minimize the limitations, to a large extent by using virtual interview methods.
8. Is the sampling strategy described?	Yes	In order to select places to visit for data collection, a purposive sample was used which included careful stratification based on geography, populations characteristics and connection with UNFPA interventions. The evaluators note that it was not a representative sample but was sufficient to show both positive and negative aspects of the work done. Because data on specific villages in terms of interventions were not always available, when the selected village did not have a UNFPA element, the next village was chosen.
9. Does the methodology enable the collection and analysis of disaggregated data?	Yes	The methodology was designed to acquire data that could be disaggregated by sex, both in documents where this was shown, and in interviews and focus groups which were designed to show gender differences. In fact, Table three shows the stakeholders who were met by institutions and gender, with 634 of the 924 stakeholder consulted being women.
10. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, disability inclusion, gender equality and human rights)?	Yes	The methodology was designed to deal with cross-cutting issues, of which there was an emphasis on both gender and disability. The evaluators noted that disability only became a focus for UNFPA in 2020 and as a result was not well-reported before that. The evaluators made an effort to obtain data about disability in terms of the areas of action but noted that this proved difficult because of the constraints on data collection caused by COVID-19. As noted in the report, "Due to COVID 19 restrictions and travel difficulties, interviewing people with disabilities was difficult. To mitigate this, NGOs representing people with disabilities were met to obtain data, but those were a very small number, and the findings cannot be generalized to a larger population. Questions were posed to midwives on whether women with disability came to access institutional delivery and to IPs whether girls and boys with disability accessed life skill programs. However, they did not come to meetings with [the evaluation team] ET."
<b>3. Reliability of Data</b>	Yes No Partial	Assessment Level: <b>Very good</b>
<i>To ensure quality of data and robust data collection processes</i>		
1. Did the evaluation triangulate data collected as appropriate?	Yes	Throughout the findings, the different sources of data are applied so that they can show clearly the results. Triangulation was discussed and the evaluation matrices (which are shown in considerable detail as annexes) also describe triangulation.
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	The data sources were discussed in detail in the text and they are shown in annexes. In the discussion of limitations, absence of data, for example on disability, was noted.
3. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Yes	The norms relating to discrimination and ethical considerations were followed in the evaluation. There was a specific discussion of how ethical considerations were addressed on page 9 which, among other things, noted that "Ensuring the protection of respondents' rights, an informed consent was sought before all interviews were made." In the presentation of data, the general source of quotations was noted, but not the name or position of the stakeholder quoted.

<b>4. Analysis and Findings</b>	Yes No Partial	Assessment Level:	<b>Very good</b>
<i>To ensure sound analysis and credible findings</i>			
<b>1. Are the findings substantiated by evidence?</b>	Yes	There were 36 findings distributed according to evaluation questions. In each case, there was an extensive presentation of the evidence supporting the finding. They were based on a combination of document review, interviews, focus groups and observations, which were referenced by footnotes in the text. The annexes, however, provided even more detail.	
2. Is the basis for interpretations carefully described?	Yes	The evaluators were careful to explain and show the basis for each finding.	
3. Is the analysis presented against the evaluation questions?	Yes	The findings were structured by criteria and within them by questions.	
<b>4. Are the cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?</b>	Yes	The evaluators described clearly the causal connections between UNFPA contributions and the observed results. This was done throughout, but was particularly notable in the discussion of gender. In addition a number of unintended outcomes were identified including, for example, "The expanded use of the e-learning platform created for capacity building of providers for the programme related to implementation of Antara programme was unintended and expanded the scope and reach of the platform."	
5. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The target groups for the UNFPA country programme are many and diverse. Throughout the findings an effort was made to show how group factors were important. This was again particularly notable in the findings on gender, where there was a variety of target groups for UNFPA support.	
6. Is the analysis presented against contextual factors?	Yes	In all of the findings, contextual factors like government structure, emergencies, funding and other issues were shown. It is also notable that each criterion covered ended with a subsection on facilitating and hindering factors.	
7. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, disability inclusion, gender equality and human rights?	Yes	While there was a general inclusion of equity and vulnerability factors, there was an extensive analysis of gender equality. As described above, disability inclusion was addressed to the extent data was available.	
<b>5. Conclusions</b>			
	Yes No Partial	Assessment Level:	<b>Very good</b>
<i>To assess the validity of conclusions</i>			
1. Do the conclusions flow clearly from the findings?	Yes	In each conclusion, the finding or findings on which it is based is shown.	
<b>2. Do the conclusions go beyond the findings and provide a thorough understanding of the underlying issues of the programme/initiative/system being evaluated and reflect as appropriate cross-cutting issues such as equality and vulnerability, disability inclusion, gender equality and human rights?</b>	Yes	The conclusions summarize sets of findings and in most cases place them in a larger context. They also suggest where there are needs or opportunities to improve performance. A good example is conclusions 20 and 21 about results monitoring. Cross-cutting issues, including disability, are covered in the conclusions, especially gender equality.	
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	There is no evidence of bias.	
<b>6. Recommendations</b>			
	Yes No Partial	Assessment Level:	<b>Good</b>
<i>To ensure the usefulness and clarity of recommendations</i>			
1. Do recommendations flow logically from conclusions?	Yes	The recommendations are structured around the conclusions (and combine a number of them in many of the recommendations).	
<b>2. Are the recommendations targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?</b>	Partial	There are 11 overall recommendations, each with an Action Plan that includes a mix of sub-recommendation and implementation advice. At the beginning of the section it is noted that the intended users are the UNFPA Country Office and that the recommendations are intended to help formulate a more effective Country Programme 10. The evaluators also note that they did not have information on the budget for CP 10 and could not suggest financial implications. Although all necessary elements are included, this section would benefit from a clearer formulation. In several cases the main recommendation could be more clearly stated and all should be better formatted so that they stand out from the other information such as priority level and associated conclusions.	
3. Do recommendations appear balanced and impartial and address, as relevant, key cross cutting issues such as equity and vulnerability, disability-inclusion, gender equality and human rights?	Yes	The recommendations are impartial and include gender equity, human rights and disability inclusion.	
4. Are the recommendations prioritized?	Yes	Each recommendation is prioritized as either high or medium priority.	

7. Gender	0 1 2 3 (**)	Assessment Level:	Very 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?			<p><b>a. Does the evaluation include an objective specific to assessment of human rights and gender equality considerations or was it mainstreamed in other objectives? (Score: 0-3)</b> There was no specific gender objective but it was included as part of the evaluation scope. Score 3</p> <p><b>b. Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria? (Score: 0-3)</b> There was a stand-alone gender evaluation criterion. Score 3</p> <p><b>c. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation? (Score: 0-3)</b> There were two stand-alone evaluation questions. Score 3</p> <p><b>d. Does the evaluation assess whether sufficient information was collected during the implementation period on specific result indicators to measure progress on human rights and gender equality results?(Score: 0-3)</b> The evaluation was careful to assess whether there was sufficient information, and notes cases where there was not, although mostly there was enough. Score 3</p>	
2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?			<p><b>a. Does the evaluation specify how gender issues are addressed in the methodology, including: how data collection and analysis methods integrate gender considerations and ensure data collected is disaggregated by sex? (Score: 0-3)</b> The evaluation highlights the integration of a HR and GE approach and briefly discusses how gender issues will be measured, including gender disaggregation. It is notable that the evaluation team included a gender expert. Score 3</p> <p><b>b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEEW considerations (collecting and analyzing both quantitative and qualitative data, and ensuring the appropriate sample size)? (Score: 0-3)</b> There is a mixed method approach and over 2/3 of the interviews are with women and there were gender-specific focus groups. Score 3</p> <p><b>c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility? (Score: 0-3)</b> Diverse sources were used, including especially gender-based focus groups, but also use of gender questions in interview formats. Score 3</p> <p><b>d. Do the evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate? (Score: 0-3)</b> The sampling process accounted for this diversity. Score 3</p> <p><b>e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality? (Score: 0-3)</b> There is a specific sub-section that adequately covers ethical considerations. Score 3.</p>	
3. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?			<p><b>a. Does the evaluation have a background section that includes an intersectional analysis of the specific social groups affected by the issue or spell out the relevant normative instruments or policies related to human rights and gender equality? (Score: 0-3)</b> There is a background section that adequately covers gender. Score 3.</p> <p><b>b. Do the findings include data analysis that explicitly and transparently triangulates the voices of different social role groups, and/or disaggregates quantitative data, where applicable? (Score: 0-3)</b> There were four specific findings on gender, but other findings also dealt with gender roles and were based on disaggregated data. There is some use of participant quotes to highlight the perspectives of different stakeholder groups. Score 3</p> <p><b>c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3)</b> There were a number of unanticipated effects, especially relating to a midwifery program, that were described. Score 3</p> <p><b>d. Does the evaluation report provide specific recommendations addressing GEEW issues, and priorities for action to improve GEEW or the intervention or future initiatives in this area? (Score: 0-3)</b> A specific programmatic recommendation on gender was included as a priority for the next country programme. Score 3</p>	
<p>(*) This assessment criteria is fully based on the UN-SWAP Scoring Tool. Each sub-criteria shall be equally weighted.  (**) Scoring uses a four point scale (0-3).</p>				
<b>Overall Evaluation Quality Assessment</b>				
		<b>Assessment Levels (*)</b>		
<b>Quality assessment criteria (scoring points*)</b>	<b>Very good</b>	<b>Good</b>	<b>Fair</b>	<b>Unsatisfactory</b>
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>82</b>	<b>11</b>	<b>7</b>	

Overall assessment level of evaluation report **Very good**

The evaluation integrates adequately cross cutting issues of gender equality, human rights and disability inclusion, even though is not included as part of the evaluation objective.

**Consideration of significant constraints (e.g. COVID-19 or civil unrest)**

The quality of this evaluation report has been hampered by exceptionally difficult circumstances:  Yes  No

If yes, please explain: