Organizational unit:				Year of report:	2019		
Title of evaluation report:	GRZ/UNFPA 8th Country Programme Evaluation: Zambia Period covered: 2016 - mid 2019						
Overall quality of report:	Very Good			Date of assessment:	March 2020		
Overall comments:	of data, analysis and findings, conclusions, recommendations, and integration of gender analysis and triangulation. The report provides findings that reflect a good analysis of t	The evaluation he programme's ective on both t	n uses a mixed-methods a s strengths and weakness the positive and negative	pproach, well-constructed theory of chan es from a human rights and gender equalit	ons within the scope, methodology, findings and recommendations. The strengths of the evaluation were in the reliability rge, consults a diverse sample of stakeholders, adequately mitigates the limitations in data collection, and shows good data ty (GEEW) lens. The conclusions and recommendations are coherently linked and are clearly actionable, prioritized, and or the design of the next country programme. However, more attention could be given to strengthening the evaluation		
Assessment Levels	Very Good strong, above average, best Good satisfactory, respectable	Fair	with some weaknesses, s acceptable	till Unsatisfactory weak, do	bes not meet minimal quality standards		
Quality Assessment Criteria		Insert assessment level followed by main comments. (use 'shading' function to give cells corresponding colour)					
I. Structure and Clar	ty of Reporting	Yes No Partial		Assessment Level:	Fair		
	prehensive and user-friendly						
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?		Yes	The report is accessible	and easy to read and understand.			
<ol> <li>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)</li> </ol>			The report meets the 70-page standard for country programme evaluations.				
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)?			The report is structure is coherent. The use of evaluation criteria to report key evaluation findings covers the key evaluative areas in the terms of reference (ToR), including sections on country context, UNFPA Country Programme, Findings, Conclusions, Recommendations, and Lessons Learned. The authors used text boxes to provide summaries of key evaluation findings.				
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?			Annexes contain all the required information; the evaluation ToR, list of persons met, documents read, Atlas projects, evaluation matrix, data collection tools, documented achievements in Zambia against SDGs, and stakeholder analysis map and description of the stakeholder consultation process.				
Executive summary			<b>T</b>	· · · · · · · · · · · · · · · · · · ·	an American Islam and All Read Islam the state of the set		
5. Is an executive summary included in the report, written as a stand-alone section and presenting the main results of the evaluation?		Partial	The executive summary	is concise and informative however it is n	ot written as a stand-alone summary suitable for people who might not read the main report.		
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)?		Partial	Although the executive summary contains relevant information, it is not written as a stand-alone summary suitable for people who might not read the main report. For example, it lacks key sections such as the descriptions of the contextual information and the overview of the intervention.				
7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?			The executive summary is within required page limits.				
2. Design and Method	ology	Yes No Partial		Assessment Level:	Fair		
To ensure that the evaluation	To ensure that the evaluation is put within its context						

<ol> <li>Does the evaluation describe the target audience for the evaluation?</li> </ol>		The target audience was briefly described in the executive summary as the Zambia UNFPA CO, partners and stakeholders, the regional office and headquarters, and the UNFPA Execut Board. Although the ToR also indicated that other primary users of the evaluation are the Government of Zambia, Cooperating Partners and the Implementing Partners. This however not explicit in the report.		
2. Is the development and institutional context of the evaluation clearly described and constraints explained?		The evaluation report begins with a summary of the key facts sheet of Zambia which are further explained in the contextual description. The report provides relevant contextual information on the key social, political, economic, demographic, and institutional factors that have a direct bearing on the object is described. It is noted in the report that "The key challenges regarding population and development in Zambia are high fertility and population growth, insufficiently strong governance, gender inequality, poverty, and inadequate dissemination" (p. 8). In addition to the description of the development challenges and the partner government's strategies and priorities in the National Development Plan, the report als provides a description of UNFPA's Response and Programme Strategies in the 7th and 8th Country Programme. Finally, the report provides country progress against SDGs 3, 4 and 5 relevant to the UNFPA mandate in Table 1.		
3. Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change, and assess the adequacy of these?		The evaluation report provides an elaborate theory of change in Figure 1 on page 3, and descriptions of the reconstruction of the ToC compared to the original. Although the evaluators discuss the changes and use of the ToC, description of what the conceptual and logical linkages or what the arrows between outputs, results, outcomes, the assumptions and risks represent would be useful for the reader to understand how change is conceptualized for the intervention.		
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?		The evaluation framework is not fully described in text, but the qu the evaluation questions, assumptions, indicators, data sources an	estions are presented on page 4 and in the elaborate evaluation matrix in Annex 5 (pp. 99 - 103). The matrix includes d methods for data collection.	
5. Are the tools for data collection described and their choice justified?	Yes	The data collections tools are described and the interview protocols for different stakeholder groups are provided in Annex 6. Thorough justification is provided for site visits, but is le explicit for the KIIs and group discussions. This section would be stronger if the application of the tools was further explained, such as number site visits/interviews/focus groups, and number of respondents for each, with disaggregation by stakeholder groups and gender as relevant.		
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 Stakeholder map and description of stakeholder consultation process is provided in Annex 8 (p. 111 - 114) and also briefly in the methodology section of the report. It is noted in the report that, "The evaluation team undertook further analysis and drafting of a zero draft evaluation report for review by the CO. The revised report was submitted to the ERG and further validated through a key stakeholders' meeting during which consultations on draft recommendation were undertaken" (p. 6).		
7. Are the methods for analysis clearly described for all types of data?	Yes	The report provides a brief description of the methods of data analysis and triangulation. The range of techniques for data analysis included content analysis, contribution analysis, and trend analysis.		
<ol> <li>Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)</li> </ol>		The methodological limitations are clearly articulated, i.e. little time for data collection, premature withdrawal of the national consultant for adolescents and youth, and loss of some data fieldwork data. Actions to mitigate and reduce these limitations are also provided.		
9. Is the sampling strategy described?	Yes	The evaluation team used purposive sampling approach to select sites for visits and stakeholders for interviews and focus group discussions. The criteria for selection were: stronger weaker IPs, and financially large and small programmes and projects; partners from government and civil society organisations (CSOs), donors and, importantly, primary and secondar beneficiaries.		
10. Does the methodology enable the collection and analysis of disaggregated data?	Yes	The report is explicit about how the methodology will enable dat.	collection and analysis of disaggregated data.	
II. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Yes	Yes, the evaluation framework explicitly show how cross-cutting issues (gender, human rights, and disability responsive) were assessed. The evaluation matrix includes one question and one assumption and indicators for the effectiveness criteria specifically to address cross-cutting issues. Furthermore, it was noted in the report that "Primary beneficiaries included wome girls and boys (interviewed as separate groups) reached with UNFPA-supported interventions the list of people consulted include representatives of members of youth groups in focus groups discussions" (p. 5).		
3. Reliability of Data	Yes			
	No	Assessment Level:	Very good	
	Partial	<u> </u>		
To ensure quality of data and robust data collection processes are used in the findings				
I. Did the evaluation triangulate data collected as appropriate?	Yes		ata to validate the findings from the qualitative data collected from key informant interviews, focus groups, and site rided a description of how they used triangulation validating qualitative findings. It is not clear how triangulation was done.	

	No Partial	Assessment Level:	Very good		
5. Conclusions	Yes				
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes	The analysis includes extensive coverage of cross-cutting issues, including findings which focus on the harms to girls caused by child marriage, the needs and rights of adolescent girls and boys to access SRH information and services, the SRHR of people with disabilities, the empowerment of women to address unequal and harmful power dynamics, and gender sensitive family planning, including male and female condoms, and empowering women and young people to engage in policy dialogues on rights. It was noted that the needs and rights of key populations of sex workers, MSM and lesbian, gay, bisexual, transgender and intersex (LGBTI) populations are insufficiently addressed.			
7. Is the analysis presented against contextual factors?	Yes	Contextual factors are reflected in the analysis and reporting. For example, it is noted that UNFPA engaged the chiefs and headmen and also religious leaders at the community level to address negative gender and social norms and unequal power relations that impede equitable access to SRH and HIV services (RF). The reported noted that UNFPA engaged the Nation Development Coordination Committee and the Cluster Advisory Groups at the national level and the Provincial Development Coordination Committees at the provincial level to promote multi-sectoral approach.			
6. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The evaluation findings presented different outcomes for different target groups, especially in the four tables that summarize the key findings per program outputs for women, adoles and youth, and sectors. The findings and analysis describe outcomes and ways that the intervention accommodated the needs of different groups, such as youth and adolescents, wor and persons with disabilities. In addition, the report consistently describes differences in the achievement of outcomes across the groups.			
5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	As noted earlier, the evaluation team presented a reconstructed theory of change model which they used throughout the report to explain the cause and effect links between the intervention and end results. The report notes some unintended negative consequences in addressing social norms and cultural practices around SRH. It was reported that delay in v accessing obstetric services is a result of behaviors of some health care personnel. It was note that, "An unintended consequence, however, was that some facility staff reported in interviews 'chasing' women who arrive for antenatal services without their partner/husband, or giving priority to those who are accompanied. As it was not possible to assess the exit to which this may actually be occurring, this is considered a tentative finding requiring further investigation" (p. 30).			
4. Is the analysis transparent about the sources and quality of data?	Yes	As noted earlier, the evaluation team referenced sources in foot provide remarks or judgment on the quality of data.	notes and in-text, although use of direct quotes could improve the transparency of the analysis. The report does not		
3. Is the analysis presented against the evaluation questions?	Yes	The relevant evaluation questions are listed at the beginning of the discussion of each criteria, and the analysis is structured accordingly.			
<ol> <li>Is the basis for interpretations carefully described?</li> </ol>	Yes	This is well done throughout the Findings section, particularly evident in the "Effectiveness" section where baseline and end-of-project data is compared. Furthermore, the evaluators the theory of change in drawing conclusions and recommendations. It is noted that "This involved exploring the theory of change in the results chain logic for each component area of country programme. The linked trend analysis explored the change in results over time in the quantitative indicators of the CP, leading to conclusions and recommendations concent the appropriateness and sufficiency of indicators, outputs and targets, and noting factors that may have made effective monitoring a challenge" (p. 5).			
<ol> <li>Are the findings substantiated by evidence?</li> </ol>	Yes	summarizing key findings per question in textboxes and providin program outcomes, outputs, indicators and shows achievements	t the report showed that the evaluators substantiated findings with evidence. The evaluation team did a good job g evidence in tables (7, 8, 9, 10, and 11) in addition to detailed explanation in text. The tables provided information on the a gainst baselines and targets by year. Nevertheless, the team could have used more direct quotes from the qualitative da report. In fact, only one direct quote was used in chapter on Findings.		
To ensure sound analysis and credible findings					
4. Analysis and Findings	Yes No Partial	Assessment Level:	Very good		
considerations?	Yes	Throughout, the evaluation team was objective and impartial, ensuring informant confidentiality" (p. 5).			
sources and if relevant, explained what was done to minimize such issues?	Yes	data with objective qualifications, or to omit them and explain the data gap. It is stated if robust conclusions and recommendations on a specific area could not be drawn" (p. 5).			
<ol> <li>Did the evaluation make explicit any possible limitations (bias, data gaps etc.) in primary and secondary data</li> </ol>	Yes		bivalent or conflicted, the evaluation team undertook further interviews, phone contact and/or document review" (p. 5).		
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?		The evaluators were explicit about the source of qualitative data coming mainly from interviews, focus groups, and observations during site visits. The evaluators cauti consultant did not have full access to field notes when the national consultant for adolescents and youth was withdrawn from the evaluation. However, the lead consul more interviews to address this concern. Furthermore, it is noted in the report that evaluators assessed the change in results over time using the quantitative indicato			

To assess the validity of conclusions					
1. Do the conclusions flow clearly from the findings?		The conclusions are clearly drawn from the findings. The evaluate	ors have specified the respective evaluation question and evaluation criteria that are linked to each conclusion.		
	Yes				
	103				
2. Do the conclusions go beyond the findings and provide a thorough understanding of the underlying issues of			I programmatic level conclusions covering all programmatic components. The categorization of the conclusions into		
the programme/initiative/system being evaluated?			iderstanding of the country and program contexts and system being evaluated. Nevertheless, this evaluation adopted a : to which the observed differences/results were a consequence of the 2016-2020 CP implementation. This analytic		
	Yes		perational environment of program implementation, but not enough to understand underlying sociocultural and		
		sociopolitical contexts of the CP. A systems theoretical orientati			
3. Do the conclusions appear to convey the evaluators' unbiased judgement?		The conclusions are clearly based on evidence from the findings a	and therefore do not convey his:		
5. Do the conclusions appear to convey the evaluators unbiased judgement:	Yes	The conclusions are clearly based on evidence iron the infinities a	and therefore do not convey bias.		
6. Recommendations	Yes				
	No	Assessment Level:	Very good		
	Partial				
To ensure the usefulness and clarity of recommendations					
I. Do recommendations flow logically from conclusions?			ized by the evaluation questions, conclusion, and classified into high or medium priority of both the strategic and		
	Yes	programmatic recommendations.			
2. Are the recommendations clearly written, targeted at the intended users and action-oriented			ating 'operational implications' and defining the intended users (which also align with the intended audience of the		
(with information on their human, financial and technical implications)?		evaluation). The technical implications of the recommendations	are implicit and evaluators made general references to the human and financial implications.		
	Yes				
3. Do recommendations appear balanced and impartial?		The recommendations flow clearly from the findings and conclusi	ions without any evidence of bias.		
	Yes				
		The account of a size and a size of			
4. Is a timeframe for implementation proposed?		of the 8CP in 2020 and the next programming cycle (9CP) in 202	n a 1-2-year period. It was noted the actionable recommendations will largely inform 2020 annual planning, the extension		
	Yes				
5. Are the recommendations prioritized and clearly presented to facilitate appropriate management response		The classification into high and medium seems appropriate to fac	ilitate management response, considering that the evaluators got feedback from the CP stakeholders on the		
and follow up on each specific recommendation?	Yes	recommendations.			
	Tes				
7. Gender	0				
	I	Assessment Level:	Very good		
	2				
To means the interaction of Condex Faultic and Factorization of Manage (CFFMA (#)	3 (**)				
To assess the integration of Gender Equality and Empowerment of Women (GEEW) (*)					
<ol> <li>Is GEEW integrated in the evaluation scope of analysis and indicators designed in a way that ensures GEEW- related data to be collected?</li> </ol>		(Score: 0-3)	essment of human rights and gender equality considerations or was it mainstreamed in other objectives?		
			onsideration in the evaluation scope of analysis with clear indicators and specific assumptions against which data was		
			t extent has the programme mainstreamed gender and human rights-based approaches including for people with		
		disabilities?" (Score = 3).			
			ghts included in the evaluation framework or mainstreamed into other evaluation criteria? (Score: 0-3)		
			with clear assumption and three indicators is included within the Effectiveness criterion is included in the evaluation matrix the Relevance criterion. The report noted that CP8 is contributing to relevant policies especially on gender although more		
			policy level. The evaluation shows that UNFPA is contributing to relevant laws and policies to address factors affecting		
			d violations of human rights (especially in sub-section 4.1.1.2) (Score = 3).		
			on regarding how GEEW was integrated into the subject of the evaluation? (Score: 0-3)		
			e strategic gender needs through policy interventions and EQ2 to address practical gender needs through effective service		
		delivery (sub-sections 4.2.1, 4.2.2 and 4.2.3) (Score = 3).			
		d. Does the evaluation assess whether sufficient informat	ion was collected during the implementation period on specific result indicators to measure progress on		
		human rights and gender equality results? (Score: 0-3)	and a specific result indicators to measure provide on specific result indicators to measure progress on		
			ency). For example, it is noted that "During the 8CP UNFPA played a significant role in enhancing capacity in coordination		
		and harmonization in the generation of statistical data in Zambia	through the National Strategy for the Development of Statistics (NSDS)" (p. 41). Furthermore, it is noted in the report		
		с <i>,</i>	further strengthening" (p. 47) and that "Most IPs do not have dedicated, trained staff for M&E, and the quality of M&E		
		among IPs needs to improve, particularly for government (IP repo	orts, KI interviews) (p. 49) (Score = 3).		

2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?		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) The evaluation methodology was gender responsive with regards to the methods and tools for data collection. For example, it is noted in the report that key informant interviews included women, girls and boys (interviewed as separate groups) as the primary beneficiaries of UNFPA-supported interventions. However, the report does not provide a gender disaggregated list of persons consulted (Score = 3).
		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) There is no description in the methodology section to show collection and analysis of both qualitative and quantitative data was done to ensure appropriate sample size for evaluating GEEW considerations. However, it is noted in narrative in Annex 8 that the evaluation team used purposive sampling from the stakeholder map to get samples for FGDs and KIIs. It is not clear if the team used a gender, responsive stakeholder analysis to determine the numbers of primary beneficiaries or rights holders (women, adolescents and youth, particularly girls) and secondary beneficiaries or duty bearers (implementing partners and government staff such as health staff trained with UNFPA support) when selecting samples. (Score = 2).
	3	c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility? (Score: 0-3) Yes, it is noted in the Executive Summary that "[t]he evaluation team undertook data capture, analysis and triangulation utilizing extensive document review, field visits, semi-structured key informant interviews in the office and with stakeholders drawn from the full range of partners, and focus group discussions with primary and secondary beneficiaries. The evaluation was thus highly participatory and full confidentiality was assured" (p. xi). The team got feedback from Co and ERG on both the draft and final reports to validate accuracy of findings, thus ensured the credibility of the data. (Score = 3)
		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) Yes, this was ensured with samples representing the primary beneficiaries as noted above. (Score = 3).
		e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality? (Score: 0-3) The interview and focus groups protocols indicate that data collectors complied to ethical guidelines for informed consent and respect for confidentiality. Although the report is not explicit about the appropriateness of interview and focus groups locations, it is noted in the report that "Primary beneficiaries included women, girls and boys (interviewed as separate groups) reached with UNFPA-supported interventions" (p. 5). (Score = 3).
3. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?		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) The evaluation report has a background section that includes an alignment to the national policies, strategies, action-plans, and relevant national stakeholders working on human rights and
		gender equality. It is noted in the report that "Notably the CP is aligned to the Anti-Gender Based Violence Act of 2011, the GBV National Guidelines of 2011, and the 2014 National Gender Policy. The 8CP is actively engaged in integrated SRH, including for young people, within these policies and frameworks" (p. 22). (Score = 3).
		<ul> <li>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)</li> <li>While the findings include data analysis that explicitly triangulates the voices of different groups, this is not consistently done. Quantitative data is used especially in the tables that show achievements against baselines and targets, but there is, for example, minimal use of quotes or stories to highlight the voices of different groups or clear gender-disaggregated data. (Score = 2).</li> </ul>
	3	c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3) As noted above in Question 4.5 under Key Findings section, the unanticipated effects of the intervention on human rights and gender equality is reported (Score = 3).
		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) The report provides seven specific conclusions and recommendations that explicitly addresses gender equality, particularly: train Safe Motherhood Action Groups (SMAGs) and community based distributors (CBDs) (Conclusion 7, Recommendation 7); engagement with the LGBTI populations and sex workers (Conclusion 8, Recommendation 8); expansion and/or modification of the UN Joint Disability Inclusion Project (Conclusion 9, Recommendation 9); finalising the road map for fistula repair (Conclusion 10, Recommendation 10); Integrated SRHR/HIV/GBV programming for young people (Conclusion 11, Recommendation 11); intensify and expand the focus on child marriage in the next CP (Conclusion 12, Recommendation 12); and increased integration of AY, SRHR and the gender focus within the CO (Conclusion 13, Recommendation 13). (Score = 3).

(\*) 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 totaling 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.

I = 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.

				Assessment Levels (*)
Quality assessment criteria (scoring points*)	Very good	Good	Fair	Unsatisfactory
	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	<u> </u>		
I. Structure and clarity of reporting, including executive summary (7)			7	
2. Design and methodology (13)			13	
3. Reliability of data (11)	П			
4. Analysis and findings (40)	40			
5. Conclusions (11)				
6. Recommendations (11)	П			
7. Integration of gender (7)	7			
Total scoring points	80		20	
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
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				
Consideration of significant constraints The quality of this evaluation report has been hampered by exceptionally difficult circumstances: f yes, please explain:				