

GEROS Evaluation Quality Assurance Tool

Version: September 2021

Formative Evaluation of the Federal Government of Nigeria–UNICEF Nutrition Country Programme Component of Cooperation, 2018–2022

REPORT RATING SUMMARY

Overall Rating	68%	Satisfactory
●●●●● Exceptional (96% - 100%)	5	
●●●● Highly Satisfactory (87.5% - 95.99%)	4	
●●●● Satisfactory (62.5% - 87.49%)	3	Meets UNICEF/UNEG standards for evaluation reports and decision makers may use the evaluation with confidence
●●● Fair (35% - 62.49%)	2	
●● Unsatisfactory (0% - 34.99%)	1	

REPORT DETAILS

Title of the evaluation report	Formative Evaluation of the Federal Government of Nigeria–UNICEF Nutrition Country Programme Component of Cooperation, 2018–2022
Report sequence number	Nigeria-61-2022-17807
Region	WCAR
Year of report	2022
Office	Nigeria Country Office
Coverage (countries)	Nigeria
ToRs present	Yes
Date of review (dd/mmm/yyyy)	February 4, 2023
Name of review firm	IOD PARC

CLASSIFICATION OF EVALUATION REPORT

Management of evaluation (*Managerial control and oversight of evaluation*) Jointly managed with Country

Unicef goal areas (*Alignment with strategic plan priorities*)

Every child survives and thrives	Yes
Every child learns	No
Every child is protected from violence and exploitation	No
Every child lives in a safe and clean environment	No
Every child has an equitable chance in life	No
Gender equality (cross-cutting)	Yes
Humanitarian action (cross-cutting)	Yes

Evaluation object	Joint Programme
Evaluation type	Summative and formative
Evaluation strategy	Mixed methods
Evaluation design (<i>primary method used</i>)	Quasi-experimental
Evaluation level	Output & Outcome
Geographic scope	National
Primary SDG(s) covered (<i>number</i>)	2

EQA Summary: *The rater will provide top line issues for this evaluation relevant for feedback to senior management (positive and negative), summarizing how the evaluation report meets or fails to meet all criteria. As relevant, the rater will highlight best practice/added value elements and the level of complexity of the evaluation.*

On the one hand this is a very comprehensive evaluation, in both scope and implementation, and is well-designed, with all the appearance of being well-executed by the evaluation team. The context and background to the issues of mal/under-nutrition in Nigeria are well documented, as is the background and justification to this programme. Undoubtedly a considerable body of good evidence has been gathered to conduct analysis and derive findings. Data collection processes were well implemented (though lacking in some detail for the quantitative data collection), and there is a good balance of data from primary/secondary/quantitative/qualitative sources, facilitating robust triangulation. The report falls down, however, on a few key areas. Firstly, it suffers from a range of obvious formatting and proofing issues that should never have been allowed to make it into a final draft. Secondly, the organisation of the analysis and findings does not live up to what is aspired to in the design description - the analysis focuses primarily on the achievements of the various field level interventions, without clearly differentiating the policy analysis that it sets out to describe. The element of humanitarian assistance receives almost no attention. Thirdly, the findings are overly positive with respect to the CPC programme - initial impressions of strong achievement are not matched by the analysis of data, which paints a picture of much more modest achievements - this should have been more systematically and clearly presented, so stakeholders can clearly understand what they have achieved and what should be done to maximise any future programme performance.

Recommendations for Improvement: *The rater will identify topline recommendations to improve the evaluation, and be specific to the sections of the report where shortcomings were found. As relevant, resources will be cited to assist evaluation managers in overseeing future evaluations.*

Firstly, this evaluation has both formative and summative components, so should be categorised as such (as opposed to formative only, which it claims to be). Some additional explanation of how the refined theory of change was derived from the original TOC would be useful, along with some analysis regarding the quality/appropriateness of the original TOC and the rationale behind the changes. The methods, while sound, are presented in a somewhat unclear manner. The report would benefit from some review of what is being presented in the methods section and potentially more detail in Table 6, specifically distinguishing between primary and secondary quantitative and qualitative datasets. The report does not have a standalone Limitations section - with an evaluation of this complexity and size, there were likely a variety of theoretical and operational challenges, which should have been documented. Findings are not systematically laid out across the two evaluation strands noted in the design - they focus on more general implementation of activities. Conclusions and Recommendations are overly general in places, with insufficient focus on actionability and/or priority or timelines for implementation. The report requires a thorough proof-reading to pick up on a variety of textual, spelling, bookmark and formatting errors - this should have been done before the final version was approved. Some of the errors are obvious, so this was clearly neglected.

SECTION RATINGS

SECTION A: EXECUTIVE SUMMARY (weight 5%)	67%	Comments on Rating
---	-----	--------------------

Question 1. Can the executive summary inform decision-making?

	i	Is clearly presented, serves as a standalone document useful for informing decision making, and is of relevant conciseness and depth for key users (Maximum of 5 pages unless otherwise specified in ToR).	Partially	The summary is 12 pages, which is very long. It could be shortened by combining the intro and context to one page, one for the purpose and methods, compressing some of the findings, bringing lessons learned and conclusions to one page, recommendations on one page and some judicious formatting.
	ii	Includes all necessary elements (overview of the intervention, evaluation purpose, objectives and intended audience, evaluation methodology, key conclusions on findings, lessons learned if requested, key recommendations) as per the ToR.	Yes	All the elements of the main report are included in the summary.
	iii	Includes all significant information needed to understand the intervention and the evaluation AND does not introduce new information from what is presented in the rest of the report.	Partially	The summary includes a good summary of relevant information from the main report, although some elements (conclusions, lessons learned, recommendations) are simply copied and pasted from the main report and thus are not summaries and repeat the issues noted below with these section.
SECTION B:	BACKGROUND (weight 5%)		93%	Comments on Rating
Question 2.	Is the object of the evaluation clearly described?			
	i	Clear and relevant description of the intervention, including: location(s), timelines, cost/budget, and implementation status.	Yes	The intervention is well described, with extensive details on all of the locations, timelines, budget etc.
	ii	Clear and relevant description of intended rightsholders (beneficiaries) and duty bearers (state and non-state actors with responsibilities regarding the intervention) by type (i.e., institutions/organizations; communities; individuals...), by geographic location(s) (i.e., urban, rural, particular neighbourhoods, town/cities, sub-regions...) and in terms of numbers reached, with disaggregation by gender, age, disability . . . (as appropriate to the purpose of the evaluation).	Partially	The background and context sections cover all of the rights holders, the locations, interventions, duty bearers etc. include maps of the proposed locations. The description of the rights-holders in terms of targeted numbers is not covered, however. It can be assumed due to the institutional capacity building nature of the programme that the populations of the states covered are targeted, but it is not explicitly noted within this section, so the reader cannot easily grasp the scale of the programme.
Question 3.	Is the context of the intervention clearly described?			
	i	Clear and relevant description of the context of the intervention (i.e. relevant policy, socio-economic, political, cultural, power/privilege, institutional, international factors) and how context relates to the implementation of the intervention.	Yes	The context section is very comprehensive (and indeed some detail could be removed to annexes - discussed below) and covers the full gamut of the policy, socio-economic and institutional context. One small issue is in relation to citing of very recent (2021/22) statistics on nutrition in the context/background section (Section 2). The programme under evaluation started in 2018, so stats from this time would have been more appropriate to establish the requirement for the interventions more appropriately. Other stats (e.g. 2.2, pg. 3, para 3) have no reference.
	ii	Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Yes	The programme description (Section 3) clearly links the overall country programme and the nutrition subcomponent within the context of the SDGs and their associated description.
	iii	Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Yes	A very comprehensive description of the background to the rights-holders (specifically in relation to prevalence of nutritional deficiencies and their health outcomes) is provided in Section 2.
Question 4.	Are key stakeholders, their relationships and contributions clearly identified?			
	i	Identification of implementing agency(ies), development partners, right holders, and additional duty bearers and other stakeholders; and of linkages between them (e.g., stakeholder map) (if relevant).	Yes	Covered in Section 3 - Table 2 (pg. 8) provides a useful mapping of the different elements of the programme and the stakeholders. A very minor issue is that the 3 tiers to which states are allocated are not explained. The reader unfamiliar with the Nigeria context may not be clear what the basis is.
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Yes	Also covered adequately in Section 3. Table 3 (pg. 11, Section 3.3) lists the main stakeholders and their specific roles.
SECTION C:	EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)		83%	Comments on Rating
Question 5.	Is the purpose of the evaluation clearly described?			
	i	Purpose of evaluation is clearly defined, including why it was needed at that point in time, its intended use, and key intended users.	Yes	Very well laid out in Section 4.1, with a table of stakeholders (table 4) in relation to their intended contributions to and uses of the evaluation.
Question 6.	Are the objectives and scope of the evaluation clear and realistic?			
	i	Clear and complete description of what the evaluation seeks to achieve by the end of the process with reference to any changes made to the objectives included in the ToR (if applicable).	Yes	The objectives are well formulated, in line with the DAC evaluation criteria (although not fully articulated - this is done comprehensively in the evaluation matrix) and well-articulate what the evaluation seeks to achieve. No deviations from the TOR were undertaken - this is also noted explicitly.
	ii	Clear and relevant description of the scope of the evaluation: what will and will not be covered (thematically, chronologically, geographically with key terms defined), as well as, if applicable, the reasons for this scope (e.g., specifications by the ToRs, lack of access to particular geographic areas for political or safety reasons at the time of the evaluation, lack of data/evidence on particular elements of the intervention).	Partially	Sections 4.3-4.5 cover a description of the thematic, geographical and temporal scope of the evaluation. This is reasonably concise and clear, but it doesn't fully integrate with what is covered by the description in the methods section. In the context description the nutrition component of the CPC has three elements (policy/planning, service delivery and community outreach, humanitarian relief) whereas in the evaluation methods, the programming appears to be divided into two components for separate evaluation: the multisectoral integrated interventions pilot and the nationwide nutrition programmes. This is somewhat confused and would have benefited from being more clearly spelled out.
Question 7.	Is the theory of change, results chain or logic well articulated?			

	i Clear description of the intervention's intended results, or of the parts of the results chain that are applicable to, or are being tested by, the evaluation.	Yes	The theory of change is presented in Annex B. It is detailed and comprehensive, covering all the relevant elements of programming that are to be tested. The evaluation also assesses the specific results to be achieved by the programme (i.e. quantitative targets), articulated within the finding section.
	ii Causal relationship between outputs and outcomes is presented in narrative and/or graphic form (e.g., results chain, logic model, theory of change, evaluation matrix).	Yes	As above, the TOC in Annex B is quite clear in articulating the results chain of the programme, with all elements well set out. The evaluation matrix has the intended outputs/outcomes of the evaluation laid out clearly.
	iii For theory-based evaluations, the theory of change or results framework is assessed, and if requested in the ToR, it is reformulated/improved by the evaluators.	Partially	There is some lack of clarity as to the TOC that was developed for the evaluation (presented in Annex B) vs. the TOC that governed programme implementation (presented in the TORs, also annexed, pg. 114). There are clear differences between the two TOCs in terms of how they are articulated (although many overlaps also). An explanation of how the "new" TOC was derived from the "original" TOC is not provided, nor is either TOC formally evaluated or assessed for quality within the report.
SECTION D:	EVALUATION DESIGN AND METHODOLOGY (weight 20%)	82%	Comments on Rating
Question 8.	Does the evaluation use questions and the relevant evaluation criteria that are explicitly justified as appropriate for the purpose of the evaluation? <i>UNICEF evaluation standards refer to the OECD/DAC criteria - Relevance; Coherence; Effectiveness; Efficiency; Sustainability; Impact (not all are necessarily relevant for all evaluations). Evaluations should also consider equity and leaving no-one behind, gender and human rights based approach (these can be mainstreamed into other criteria). Humanitarian evaluations should also consider Coverage; Connectedness; Coordination; Protection; Security.</i>		
	i Evaluation questions and sub-questions are appropriate for meeting the objectives and purpose of the evaluation. The relevant criteria are specified and are aligned with the questions.	Yes	A comprehensive evaluation matrix has been presented in Annex C which articulates the evaluation questions. The EQs are clear, concise and well-presented with some sub-text after each to further ensure their alignment with the evaluation criteria is well-understood. Very good.
	ii In addition to the questions and sub-questions, the evaluation matrix includes indicators, benchmarks, assumptions and/or other processes from which the analysis can be based and conclusions drawn.	Yes	There are 14 evaluation questions, with no sub-questions presented - this is appropriate given the number of EQs. A range of indicators have been associated with the EQs in the matrix, cross-referenced with data collection tools and analytical methods. The relevant DAC evaluation criteria (along with Resilience and Equity/Gender Equality) are aligned with the EQs in the matrix.
Question 9.	Does the report specify adequate methods for data collection, analysis, and sampling?		
	i Evaluation design and set of methods are relevant and adequately robust for the evaluation's purpose, objectives and scope; and are fully and clearly described.	Partially	While the methods as presented are quite robust and a good mix that should lead to good data that can answer the evaluation questions, the description is somewhat unclear and suggests some confusion on the part of the authors. For example, there is no distinguishing between secondary quantitative and qualitative data - though both types are used. As primary data is distinguished in this manner the reviewer found it challenging to parse exactly what data was being collected for which intervention set. In line with the above comment on lack of clarity between the programme description and the interventions under evaluation, there also appears to be some inconsistencies between how the interventions are described - specifically, the nationwide nutrition programmes eval is noted in the first sentence in Section 6 to be those implemented between 2014 and 2020, whereas in the first sentence of Section 6.1, the text notes the nationwide nutrition programming desk review "synthesise[s] the evidence on the nutrition strategies since 2018". The list of docs/datasets in Table 6 underneath then list publications from 2014 onwards. The application of the data collection tools is an area that receives limited attention also - some more detail on the specifics of the data collection process for the household surveys (e.g. the approach to collection of the anthropometric data - equipment, exact methods etc.) would have been useful in an annex.
	ii Data sources are appropriate - these would normally include qualitative and quantitative sources (unless otherwise specified in the ToR) - and are all clearly described.	Yes	Notwithstanding the need for clearer description of the mix of data types being used (see above), the sources are appropriate to this evaluation, for example the use of a variety of secondary quantitative data - the results framework data and the MICS datasets from different years. These are complemented well with primary data collection methods, both quantitative and qualitative. The Difference-in-Difference and coarsened exact matching approaches to impact measurement are appropriate for this analysis.

	iii	Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or, if not, provide reasons for this).	Yes	The sampling strategy for the impact evaluation component (hh-based surveys) (Section 6.3.3) is well-written and clearly laid out, including the main limitations and mitigations strategies for challenges, such as absence of household lists. The strategy for the qualitative data collection for both strands of the evaluation is similarly well described (Section 6.4.2), with both the sectoral and geographic sampling strategy noted. That said, the process for identifying key informants and discussion group attendees was not covered in this. Given the (relatively) small sample of qualitative research targets, some more explanation of this strategy was warranted.
	iv	Clear and complete description of the methods of analysis.	Partially	Curiously, while the report includes a concise yet complete description of the analysis of the qualitative data, no similar description of the treatment of the primary quantitative data is provided, nor is a reference to the annexes provided (other than referring to the data collection tools in Annex E), and a reference to use of SurveyCTO for survey data in section 7. This seems like an oversight given the high quality of the design and sampling description.
	v	Methodology allows for drawing causal connections between outputs and expected outcomes.	Yes	The methods are robust, with a mix of approaches taken with a mix of stakeholders. If correctly applied, they would be very well suited to drawing causal connections.
	vi	Clear and complete description of limitations and constraints faced by the evaluation, including gaps in the evidence that was generated and mitigation of bias, and how these were addressed by the evaluators (as feasible).	Partially	Two of the methods subsections dealing with a specific data collection approach have some discussion of limitations and mitigation strategies applied, specifically sections 6.2.3 and 6.3.4. While these are welcome and thorough, there is no overall limitations section in the report that summarises the various challenges that might have been faced.
Question 10.		Are ethical issues and considerations described? The evaluation should be guided by the UNEG ethical standards for evaluation. As such, the evaluation report should include:		
	i	Explicit and contextualized reference to the UNEG obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability) and/or the principles in the 2020 revised UNEG Ethical Guidelines (integrity, accountability, respect, beneficence).	Yes	The report has a comprehensive approach to, and description of, the approaches to ethical and safe data collection in Sections 7.1-7.4. There is a clear discussion of compliance with United Nations ethical standards in Section 7.4.
	ii	Description of ethical safeguards for participants appropriate for the issues relevant to methodology and how they are applied (respect for dignity and diversity, right to self-determination, fair representation, compliance with codes for vulnerable groups, confidentiality, and avoidance of harm). For those cases where the evaluation involves interviewing children, explicit reference is made to the UNICEF procedures for Ethical Research Involving Children.	Yes	The comprehensive approach to ethical and safe data collection is well described in Section 7, and full consent processes/forms are included in the annexes and IRB approval was obtained. However, there is no explicitly reference to the UNICEF procedures Ethical Research Involving Children - neither in the report itself nor in the TORs. This is an oversight which is compensated for by the robust references to other standards.
Question 11.		Does the evaluation incorporate innovative practice that adds value to the evaluation process?		
	i	Innovation practice is used to improve the quality of evaluation process. This could be evident in several ways such as the design of the methodology (i.e. use of technology for data gathering, extensive participatory processes, systematic analysis processes such as collaborative outcomes reporting and incorporation of big data, specific strategies to address complexity such as outcome harvesting, strong child rights focus), or ways of sharing of evaluation results.	Partially	The TORs encourage the use of innovative approaches to data collections such as remote data collection, digital tools or mobile SMS - this use of digital data collection tools (SurveyCTO and tablets) for the survey data was undertaken. This is a useful approach, but is increasingly mainstream rather than innovative. No specifically innovative tools were used within the methodology as it built (for quant data at least) on earlier datasets for comparison purposes.
SECTION E:		EVALUATION FINDINGS (weight 25%)	58%	Comments on Rating
Question 12.		Do the findings clearly address all evaluation objectives and scope?		
	i	Findings marshal sufficient levels of evidence to systematically address all of the evaluation's questions, sub-questions and criteria.	Partially	The findings section is quite robust in terms of the quantity of evidence that it presents to address the evaluation questions - however, the preamble to the findings section clearly implies that there will be evaluation of two separate strands (see above), but the findings just seems to cover one strand (the nutrition interventions). The humanitarian component of the programme receives little or no attention - the findings section (8.1.1, pg. 31 contains a single paragraph, in a different font, suggesting it was pasted into the document as an afterthought).
	ii	Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Partially	There is good use of the CPC results framework data as an initial point of analysis to complement the primary data. There is more limited usage of either the original CPC theory of change or the revised one presented in the annexes - there is some discussion of inconsistencies or where assumptions did not hold, but this is quite limited.
Question 13.		Are evaluation findings derived from the conscientious, explicit and judicious use of the best available, objective, reliable and valid data and by accurate quantitative and qualitative analysis of evidence.		

	i Evaluation uses credible forms of qualitative and quantitative data. It presents both output and outcome-level data as relevant to the evaluation framework. Triangulation is evident through the use of multiple data sources.	Yes	Overall, there is a wealth of qualitative and quantitative data presented throughout the findings section, indicating a robust approach to the design and analysis. Many of the findings are well-presented within their sections, but lack triangulation between sources, even where the data is there. For example, quantitative results related to the headline nutritional indicators are presented in the impact section, but the discussion around some of the reasons for poor achievement of nutritional progress for children - discussed in other sections with reference to the qualitative data (FGDs and KIIs) are not used to explore the possible drivers of these results. In this the report findings have not been optimised.
	ii Findings are clearly supported by, and respond to, the evidence presented, including both positive and negative. Findings are based on clear performance indicators, standards, benchmarks, or other means of comparison as relevant for each question.	Partially	The evaluation uses a great deal of evidence to generate findings, with a given finding typically supported by multiple sources of evidence. Further, the use of measures like "partially achieved" for programme targets (see table 14) is questionable - for some indicators, failure to achieve the target led to a rating of "not achieved" even if some progress was noted (e.g. prevalence of stunting), whereas others with some progress were rated "partially achieved". Also, the same table notes regress on indicator 2.5 (children with minimum acceptable diet), but rates it as 'partially achieved' - I suspect a typographical error here.
	iii Causal factors (contextual, organizational, managerial, etc.) leading to achievement or non-achievement of results are clearly identified. For theory-based evaluations, findings also analyse the logical chain (progression -or not- from implementation to results).	Partially	While many causal factors are noted throughout the text, drawing on the primary evidence - often the FGDs with community members or KIIs with field-based professionals (and thus of high value), there are issues with the clarity of the presentation in many cases. For instance, presentation of impact data related to the household survey is frequently simply referred to the regression analysis results tables in the Annexes, and data from the DHS or MICS presented graphically in the main report - more attention should have been paid to seek to present all data together, to ensure clarity in the findings.
Question 14.	Does the evaluation assess and use the intervention's Results Based Management elements?		
	i Assessment of the adequacy of the intervention's monitoring system (including completeness and appropriateness of results/performance framework - including vertical and horizontal logic, M&E tools and their usage) to support decision-making.	Partially	This was not explicitly part of the evaluation design. However, monitoring data was used as a secondary data source and indeed a review of the performance of the programme using monitoring data was undertaken. Hence, some assessment of the quality of this data and the system used to collect the data would have been useful.
SECTION F:	EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 10%)	50%	Comments on Rating
Question 15.	Do the conclusions clearly present an objective overall assessment of the intervention?		
	i Conclusions are clearly formulated and reflect the purpose and objectives of the evaluation. They are sufficiently forward looking (if a formative evaluation or if the implementation is expected to continue or have additional phase).	No	Firstly, there are two conclusions sections - Section 9 and then again in Section 11. This can only be an (obvious) error and underscores the observation by this reviewer that the report was not adequately reviewed internally before publication. That said, the first conclusions section (the largest one), while well-written, clear and presenting useful data, is quite short (<1.5 pages), and thus not fully reflective of the comprehensiveness of the evaluation. They would have benefitted from being more systematically structured in accordance with the outline of the report, and linking to specific findings, with more forward-looking statements that point to recommendations and lessons. The <i>second</i> conclusions section (Section 11) is a bare paragraph, and the headline statement - "The Nutrition CPC achieved its targets across a number of expected results" - is a misrepresentation of the actuality. The CPC, in fact, achieved only 33% of its intended targets for the 12 headline indicators measured (per Table 14).
	ii Conclusions are derived appropriately from findings, and present a picture of the strengths and limitations of the intervention that adds insight and analysis beyond the findings.	Partially	The lack of detail within the conclusions section means that many of the findings are thinly presented - this gives a somewhat overly-positive perspective of the programme, which a more careful reading of the findings does not necessarily support. More detailed conclusions, linked systematically to findings throughout the report, would improve the robustness and be of greater use for future improvements and redesign.
Question 16.	Are logical and informative lessons learned identified? [N/A if lessons are not presented and not requested in ToR]		
	i Identified lessons stem logically from the findings, have wider applicability and relevance beyond the object of the evaluation.	Yes	The lessons are indeed derived from the data and findings, and are relevant beyond the evaluation object - but this is one of the weaknesses of the evaluations, insofar as the lessons are very general, and are couched in terms of recommendations rather than good (or poor) practices.

	ii Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Partially	While the lessons are clearly stated, and not overly-long, in places they are superficial and could benefit from more granularity. For example, the first lesson notes "It is clearly [sic] that many of the progress made prior to COVID time were lost during the pandemic. The efforts to bounce back to reducing malnutrition should be strengthened". This is (a) such an obvious statement that it is largely worthless, and (b) not a lesson to be learned, but a recommendation. This last point is relevant to many of the lessons - they are framed more in terms of challenges faced by the programme that should be overcome, rather than a compilation of practices that should or should not be continued.
SECTION G: RECOMMENDATIONS (weight 15%)		50%	Comments on Rating
Question 17.	Are recommendations well grounded in the evaluation?		
	i Recommendations align with the evaluation purpose, are clearly formulated and logically derived from the findings and/or conclusions.	Partially	Insofar as they stand, the recommendations are in line with the evaluation purpose, and recommend a variety of policy/strategy and programmatic approaches for the main stakeholders to implement. They suffer from the same issue as the conclusions, however, in that they are somewhat general in nature, which does not do justice to the comprehensiveness of the data that has been collected.
	ii Recommendations are useful and actionable for primary intended users and uses (relevant to the intervention); guidance is given for implementation, as appropriate.	Partially	As above, some (though not all) of the recommendations are quite general. For example, Recommendation 1 (to the Govt of Nigeria) essentially recommends to implement a policy that is already in place and being implemented. Recc 1 for UNICEF essentially recommends to support the Govt in its implementation of it's policy - which is already UNICEF's mandate. These are so general as to be useless. Some of the recommendations (e.g. #5 for Govt, many of the sub-recommendation for UNICEF) are more specific and actionable.
	iii Process for developing the recommendations is described, and includes the Involvement of duty-bearers, as well as rights holders when feasible (or explanation given for why they were not involved).	Partially	A brief description of the stakeholders involved in recommendations development is presented in the Section 12 preamble. Stakeholders were exclusively institutional, with no reference to involvement of rights-holders.
Question 18.	Are recommendations clearly presented?		
	i Clear identification of groups or duty-bearers responsible for action for each recommendation (or clearly clustered group of recommendations). Clear prioritization and/or classification of recommendations to support use.	Partially	The recommendations section is divided into three target groups of stakeholders - Govt of Nigeria, UNICEF Nigeria and "Donors". Given the wide variety of stakeholders involved in this programme (per Table 2), a more granular and targeted breakdown of recommendations was warranted, as well as prioritisation/classification (not done).
SECTION H: REPORT STRUCTURE AND PRESENTATION (weight 5%)		64%	Comments on Rating
Question 19.	Does the evaluation report include all relevant information?		
	i Opening pages include: Name of evaluated object, timeframe of the evaluation, date of report, location of evaluated object, names and/or organization(s) of the evaluator(s), name of organization commissioning the evaluation, table of contents -including, as relevant, tables, graphs, figures, annexes; list of acronyms/abbreviations, page numbers.	Yes	The overall structure of the report is good, with all of the relevant information included and relevant tables, figures, annexes etc.
	ii Annexes include: terms of reference, evaluation matrix, list of interviewees, results chain/ToC/logical framework (unless included in report body), list of site visits, data collection instruments (such as survey or interview questionnaires), list of documentary evidence. Other appropriate annexes could include: additional details on methodology, information about the evaluator(s), etc.	Yes	All these elements are present and correct.
Question 20.	Is the report logically structured?		
	i Structure is easy to identify and navigate (for instance, with numbered sections, clear titles and sub-titles, well formatted).	Partially	While the structure of the report is adequate, some of the formatting is not right, for example line spacing in the Executive Summary is off, the table of contents changes font halfway through (some different fonts have been used in the main text also) - a solid proofread and reformatting was clearly not undertaken
	ii Structure accords to UNICEF guidelines for evaluation reports: context, purpose and methodology would normally precede findings, which would normally be followed by conclusions, lessons learned and recommendations.	Partially	Mostly this is in accordance with UNICEF guideline, but there are two conclusions sections - which may be an error.
Question 21.	Is the report well presented?		
	i Report is of reasonable length; it does not exceed number of pages that may be specified in ToR.	Partially	Some of the background section detail could be excerpted to an annex to reduce size, for example a lot of the information on the enablers of nutrition programming (2.5) is quite general. Otherwise, the report is long, but this is necessary due to the wide scope of the design and data collection.

	ii Report is easy to understand (written in accessible way for intended audience) and generally free from grammar, spelling and punctuation errors.	Partially	There are quite a few spelling/proofing mistakes (e.g. pg. 5, line 1 says \$580, instead of \$580 million), also, pg. 30 (and at least five other locations) contains a bookmark error reference clearly within the text, on pg. 49 the findings table has a repeated finding. All these errors strongly suggest that the final draft of the report was not proofread.
	iii Frequent use of visual aids (such as infographics, maps, tables, figures, photos) to convey key information. These are clearly presented, labelled, and referenced in text.	Partially	There is adequate use of graphics, tables and charts. Not all tables are labelled, e.g. on pg. 61. The colour coding for achievement at the title of each section, while visually engaging, is not consistently applied - the rating for Section 8.8 (Resilience) is noted as "Yellow" whereas the graphic is in green...
SECTION I:	EVALUATION PRINCIPLES (weight 10%)	88%	Comments on Rating
Question 22.	Did the evaluation design and style consider incorporation of the UN and UNICEF's commitment to a human rights-based approach to programming, to gender equality, and to equity?		
	i Reference and use of rights-based framework, and/or CRC, CCC, CEDAW and/or other rights related benchmarks in the design of the evaluation.	Yes	Section 7.4 on compliance with United Nations ethical standards makes appropriate reference to the relevant instruments.
	ii Clear description of the level of participation of key rights holders and duty bearers in the conduct of the evaluation (for example, a reference group is established, stakeholders are involved as informants or in data gathering).	Yes	There is a good description of the involvement of the key duty-bearers in the evaluation in Table 4 (pp 12-13) as well as how rights-holders are involved in data collection.
	iii Language is empowering and inclusive, avoiding gender, heterosexual, age, cultural and religious bias, among others; use terminology of rights holders and duty bearers; data is disaggregated by marginalized group; differential results are assessed (distribution of results across different groups).	Yes	The findings make appropriate use of language, without bias or inappropriate terminology. The data analysis is gender-disaggregated where appropriate and/or possible, with a good focus on gender-related results.
	iv Evaluation assesses the extent to which the implementation of the intervention addresses child rights and Leave No-one Behind (gender and other excluded and marginalized groups). It is disability inclusive.	Partially	Insofar as the programme is designed to address the most marginalised, the evaluation is focused on these groups. The evaluation design also assessed disability status, and while results of the prevalence of disability are presented in the annexes, no further analysis on this basis is undertaken.
Question 23.	Does the evaluation meet UN SWAP evaluation performance indicators? (Note: this question will be rated according to UN SWAP standards with detail provided below)	8	
	i GEEW is integrated in the Evaluation Scope of analysis, and evaluation criteria and questions are designed in a way that ensures GEEW-related data will be collected.	Fully integrated	Gender and equity considerations are a key element of the evaluation
	ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Fully integrated	As above, this is a key element of the evaluation design
	iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Satisfactorily integrated	As above, this is a key element of the evaluation design

SWAP Rating Guidance

<p>i GEEW is integrated in the Evaluation Scope of analysis, and evaluation criteria and questions are designed in a way that ensures GEEW-related data will be collected.</p> <p>a. 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?</p> <p>b. Does the evaluation include an objective specific to assessment of human rights and gender equality considerations or was it mainstreamed in other objectives?</p> <p>c. Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria?</p> <p>d. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation?</p>
<p>ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.</p> <p>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?</p> <p>b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEWE considerations?</p> <p>c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility?</p> <p>d. Does the evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate?</p> <p>e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality?</p>
<p>iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.</p> <p>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?</p> <p>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?</p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described?</p> <p>d. Does the evaluation report provide specific recommendations addressing GEWE issues, and priorities for action to improve GEWE or the intervention or future initiatives in this area?</p>