

GEROS Evaluation Quality Assurance Tool

Version: September 2021

Real-time Evaluation of UNICEF India's COVID Response

REPORT RATING SUMMARY			
Overall Rating	55%	Fair	
●●●●●	Exceptional (96% - 100%)	5	
●●●●●	Highly Satisfactory (87.5% - 95.99%)	4	
●●●●●	Satisfactory (62.5% - 87.49%)	3	
●●●●●	Fair (35% - 62.49%)	2	Meets UNICEF/UNEG standards for evaluation reports in some regards, but not all. Decision makers may use the evaluation with caution. Substantive improvements in some areas are needed
●●●●●	Unsatisfactory (0% - 34.99%)	1	

REPORT DETAILS	
Title of the evaluation report	Real-time Evaluation of UNICEF India's COVID Response
Report sequence number	India-64-2022-17415
Region	SAR
Year of report	2022
Office	India CO
Coverage (countries)	India
ToRs present	Yes
Date of review (dd/mmm/yyyy)	05/02/2023
Name of review firm	IOD PARC

CLASSIFICATION OF EVALUATION REPORT	
Management of evaluation (Managerial control and oversight of evaluation)	UNICEF managed
Unicef goal areas (Alignment with strategic plan priorities)	
Every child survives and thrives	Yes
Every child learns	Yes
Every child is protected from violence and exploitation	Yes
Every child lives in a safe and clean environment	Yes
Every child has an equitable chance in life	No
Gender equality (cross-cutting)	Yes
Humanitarian action (cross-cutting)	No
Evaluation object	Programme
Evaluation type	Summative and formative
Evaluation strategy	Qualitative
Evaluation design (primary method used)	Other
Evaluation level	Output
Geographic scope	National
Primary SDG(s) covered (number)	1, 2, 3, 4, 5, 6, 8, 10, 11

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.

As might be expected for a Real Time Evaluation this evaluation faced challenges of operating in a rapidly evolving context, and with a UNICEF CO which was overworked and under resourced. However, taking this into account the report is considered fair, and there are some areas which if strengthened would have resulted in more robust evidence being generated, that may have been more useful for real time course correction and decision making. Key strengths and weaknesses of the evaluation are noted below:

Strengths of the evaluations

- Well described background and context.
- Close relationship with the UNICEF CO is evident, and all changes to the evaluation were agreed after consultations with the CO (at IR stage as well as throughout). This led to an understanding of the challenges faced by UNICEF India during this difficult time.
- Good description of the impacts of COVID 19, and subsequent waves of infection, and lockdown, and strong evidence to demonstrate UNICEF's coherence with Government of India.

Weaknesses of the evaluation

- The Executive Summary could have been shorter and more tailored and succinct information provided.
- The challenges of data collection are articulated, however, the decision not to pursue any method of data collection with beneficiaries is difficult to understand even given the COVID 19 context. To also exclude community level frontline workers, such as health care workers, teachers or other community stakeholders is represents a flaw in the evaluation design, especially given the acknowledged gap in existing monitoring of whether services and supplies were reaching intended (often marginalized) people and communities.
- Due to the gaps in evidence a number of evaluation questions were not able to be answered, more detail/ analysis of how this impacted on the overall purpose and utility of the evaluation would be useful to provide (in addition to the well described exclusion of certain pillars of activity at IR stage).
- The findings did not represent diverse voices, which meant that conclusions and recommendations could not do so either.
- There very little gender disaggregated data and no gendered analysis due to lack of evidence. This is acknowledged by the evaluation, but given its approach was a participatory one with a cross cutting focus on gender and equity it is concerning that this evidence was not gathered.
- While there are overarching recommendations, most of the recommendations are those shared at the time of production (i.e. during 2021-21) as each pillar evaluation was completed. The evaluation acknowledges this, but this section could be improved if it included an indication if any of the recommendations had been implemented or taken forward between 2020/21 and the time the evaluation report was written in 2022.

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.

Recommendations for improvement are listed below, by the section of the evaluation to which they relate:

Purpose, Objectives and Scope

- While these were fit for purpose at the start of the evaluation, due to significant gaps in data (as might be expected to be discovered as a real time evaluation progresses), it would have been useful to have a process of check in and review if the purpose and utility were still relevant as the evaluation progressed, not just at IR stage.

Methodology

- While it may not be challenging to include beneficiaries during an ongoing crisis response, other ways to gather community level data should be considered, for example phone calls with community level stakeholders such as health workers, teachers or community leaders, especially in responses which have specific aims to reach marginalised groups.

Findings

- The challenges faced by different groups, and by women and children are outlined, but very little evidence provided as to how the UNICEF response was supporting the right people, in the right place at the right time.

Conclusions

- Conclusions should be more than a summary of the findings and add greater analysis and context to the evaluation.

Recommendations

- Recommendations should be updated at the end of the evaluation to ensure clarity on which have been accepted or implemented, particularly when the recommendations were given to UNICEF on a stepped chronological timescale from 2020/21, while the evaluation report was written in 2022.

Evaluation principles

- There should be specific and contextualised reference to how the evaluation team meets the UNEG obligations of evaluators.
- During a crisis response it is even more important to find ways to gather data from marginalised groups, and from women, appreciating that this is challenging and staying it has not been possible is concerning, and should be a significant area of learning and improvement for all RTE or crisis evaluations going forward.

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).	No	The Executive Summary is 9 pages long and some sections have not been summarised, but copied from the evaluation report - a more succinct summary of certain sections (especially context, findings and recommendations) would help to summarise and also allowed key points from each section to be highlighted.
	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	The Executive Summary does include all necessary elements.
	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.	Yes	The Executive Summary includes all significant information and does not introduce new information.
SECTION B:	BACKGROUND (weight 5%)	86%	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	<p>The location is described at a national scale, with the intervention starting in 2020 and implementation status noted as ongoing. The cost/ budget is also provided, and broken down by pillar of activity, and by year, although it is noted that "The evaluation agency did not receive any information about the actual utilization of the proposed budget, as a result, a cost-utilization analysis to assess the efficiency of UNICEF's response plan was not undertaken as part of this evaluation" (pp6).</p> <p>The intervention is defined as being developed in line with WHO guidance, and with the Government of India and is composed of 6 pillars of UNICEF's response plan, which are clearly described in the evaluation (and in more detail in Annex 9) as being:</p> <ul style="list-style-type: none"> •Risk Communication and Community Engagement (RCCE) •Improve Infection Prevention and Control (IPC) and provide critical medical and water, sanitation and hygiene (WASH) supplies •Support the provision of continued access to essential health and nutrition services for women, children and vulnerable communities, including case management •Support Access to continuous education, social protection, child protection and Gender-Based Violence (GBV) services •Data collection and social science research for public health decision making •Coordination, technical support and operational costs

	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).	Yes	<p>The rightsholders are described as all people affected by COVID 19 in India from 2020 from 2022. There is a specific breakdown of the data which notes</p> <ul style="list-style-type: none"> •The number of children who lost access to education during lockdown •The 50% increase in the number of children calling ChildLine in the early days of the first lockdown •The decrease in the access to prenatal care experienced by pregnant women •Increase in gender based violence •Bigger impact on poor people living in poor places/ states <p>While data for those in some at risk groups is also provided such as:</p> <ul style="list-style-type: none"> •Healthcare of pregnant women and new-born children Specific impacts on elderly people •Impact on children (social protection and education) <p>Duty bearers are broken down into three categories, and are clearly described:</p> <ul style="list-style-type: none"> • Government officials at the Central, State and District levels who were working at the planning level to contextualize support services based on the needs of the people on the ground in the COVID-19 context, • Accredited Social Health Activist (ASHA) workers, Anganwadi workers, Frontline Workers (FLWs), sanitation and health workers at the community level who were involved in delivering various supplies and services, • UNICEF India stakeholders and engaged personnel. (pp9).
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	<p>The context of the intervention is reasonably well explained in terms of Government policies, and a useful figure which sets out the timeline of the pandemic and when lockdowns were imposed (pp2). Public health measures are described and the impact of COVID 19 in India is starkly described:</p> <p><i>"The pandemic has been an unprecedented wake-up call, laying bare deep inequalities and exposing precisely the failures that are addressed in the 2030 Agenda for Sustainable Development."</i> (pp2)</p> <p>The fact that efforts to recover from the impact of COVID 19 will be greater in states which already suffered from greater burden of inequality, lower education standards and other social development indices, is clearly outlined.</p>
	ii	Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Yes	As well as outlining the overall impact on achieving the Agenda 2030 goals, the evaluation notes that COVID-19 has adversely impacted the gains made over the years on SDGs more generally, including but not limited to, SDG 1-6, 8, 10, and 11.
	iii	Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Partially	<p>The beneficiaries of the UNICEF intervention included, but were not limited to:</p> <ul style="list-style-type: none"> • The vulnerable and marginalized groups such as women, and children, • Migrants, • People living in hard-to-reach pockets/ difficult terrain, and • Tribal populations in the selected States. <p>More detail on their specific needs, broken down by group would have improved this section.</p>
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).	Partially	<p>Annex 8 provides a clear description of stakeholders broken down by nature of the partnership, and identifies whether they are development partners, duty bearers, different types of stakeholders (implementing partners or those who generate evidence), financial partners, partners who give in kind resources, etc. It is noted that the list of partners is not fully inclusive or exhaustive.</p> <p>The table setting this out (Annex 8) is clear, and well defined, although it could be improved if a stakeholder map was provided, or even just an additional column in the table which notes which state (or district) the partner operates in, or whether they are national stakeholders. This is provided for some in the narrative but it is not always clear, and does not allow links between partners to be clearly identified.</p>
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Yes	The contributions and key roles of stakeholders are clearly identified, noting financial and non financial contributions in Annex 8.
SECTION C:		EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)	90%	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	<p>The purpose of the evaluation is clearly defined:</p> <ul style="list-style-type: none"> •Providing feedback to the UNICEF India Crisis Management Team (CrMT) on relevance, coverage, efficiency, and effectiveness • Identifying challenges and bottlenecks • Involving partners and stakeholders in shaping UNICEF's crisis response to ensure it is more participatory and responsive to needs on the ground. • Acting as a real-time lesson-learning exercise that adjusts and improves planning and performance, allowing for ongoing correction of the current crisis response and collecting lessons for future health emergencies. • Identify gaps in UNICEF India's ongoing evidence-gathering efforts and recommend suggestions. • Collect data for use in future evaluation/s of UNICEF's response. <p>As a Real time evaluation it was intended usage was to provide lessons-learning' activity that focused on capturing the challenges and gaps in the UNICEF response plan on a real-time basis and providing recommendations for further improvement. As well as collated learning in the form of this evaluation to inform future programming.</p> <p>Key users are clearly defined as "<i>primary users of the RTE findings were UNICEF CRMT, the core group of senior members, staff and its partners at the national and state level ... but where appropriate, relevant findings may also be shared with UNICEF's government, UN and civil society partners involved in the COVID-19 crisis response to increase learning, accountability, collaboration and more effective and targeted response.</i>" (pp7)</p>
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	<p>The evaluation clearly describes what it will achieve, both on a month to month basis (being a RTE) and by the end. There is also a section which sets out Deviations in the Objectives from the Tor following the Inception phase (pp2), for example:</p> <ul style="list-style-type: none"> •Only evaluating 4 out of the 6 pillars of UNICEF activity (pillars, 1,2,3, and 5), so not evaluating objectives relating to pillars 4 and 6 (data collection and coordination). •Not including the final lesson learning documentation as part of this evaluation •Challenges with monthly reporting on pillars and other timing issues with monthly monitoring <p>All changes were agreed with UNICEF, and the ToR had noted that due to the nature of this RTE evaluation, changes to the process and objectives may be required throughout the evaluation, however, changes after the inception phase are not clearly documented.</p>
	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).	Yes	<p>The chronology of the evaluation is well described, with each of the 4 pillars being evaluated over a different timescale:</p> <ul style="list-style-type: none"> •Pillar 1 - March 2020 - November 2020 •Pillar 2 -March 2020 - February 2021 •Pillar 5 - March 2020 - July 2021 •Pillar 3 - March 2020 - January 2022 <p>The geographic scope is also clearly outlined with a list of the 5 states which were covered, and a short description of why, along with a map, and further detail in Annex 6.</p> <p>The thematic scope is not specifically mentioned, but given the pillars are defined by theme, it is acceptable to assume that the timeline of the pillars assessment also defines the themes which will be explored. It is also noted that pillars 4 and 6 will be excluded meaning that data collection process and coordination mechanisms will not be assessed. This is noted as a deviation from the ToR.</p>
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	There is a clear description of the interventions intended results under each pillar of activity. As this a RTE it is not expected there would be more detail or a results chain available.
	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).	Partially	While there is not a result chain or ToC, there is an evaluation matrix, and an analysis of activities under different pillars of activity, which may allow for some causal relationships to be identified.
	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.	Not Rated	Given this is a Real Time Assessment there is no ToC or results chain and the evaluation notes it is not appropriate to construct one, as part of the evaluation process.
SECTION D:	EVALUATION DESIGN AND METHODOLOGY (weight 20%)	50%	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.	Partially	Evaluation questions and sub questions are thorough and were developed following a gap analysis and activity analysis of UNICEF activities under each pillar. They are appropriate to meet the objectives of the evaluation, and criteria are specified, and aligned with the questions. However, it is unclear how some of the questions in relation to reaching marginalized groups will be answered without actually speaking to any rights holders/ beneficiaries.
	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.	Partially	The evaluation matrix includes data sources and identifies which pillars will be answered by each question, there is also a narrative description, along with more detail on how each question will be tailored for each pillar in Annex 1. However, this section could be improved if there was explicit reference made to any assumptions, and if there were any benchmarks which could be used to analyse the data. Annex 7 does include an analytical framework which outlines some indicators, but their robustness is not assessed.
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	A participatory, consultative and utilization-focused qualitative approach is described, all of which are appropriate and relevant. There is a useful diagram (which is unlabelled) on page10 which sets out the evaluation process for each pillar, and which clearly explains the methodological process. While the approach and the methodology is all relevant, and well described, there appears to be a significant gap in the design , which given this a RTE and relies primarily on qualitative evidence (interviews and FGDS) to gather data is concerning: •There is no method of data collection which will gather the views of rightsholders/ beneficiaries. The evaluation states that: " <i>Beneficiaries and community-level stakeholders were not included as stakeholders for this evaluation because (a) given the rapid nature, (b) difficulties in adjusting to the virtual modes of data collection, especially for the respondents and (c) the ongoing delay in the overall timeline due to changes in scope, it was mutually decided by the evaluation team and UNICEF R&E team not to undertake beneficiary level interviews/FGDs.</i> " (pp14) While the decision not to include beneficiaries or community level stakeholders is explained, it does represent a significant gap in the data generated, which does impact on the robustness of the methods, and of the overarching participatory approach, which cannot be completely mitigated against.
	ii Data sources are appropriate - these would normally include qualitative and quantitative sources (unless otherwise specified in the ToR) - and are all clearly described.	Partially	The primary data collection mainly relied on qualitative methods (KIIS) with different sets of stakeholders • UNICEF officials at the National and State levels, State and District/Block •Government officials and Civil Society Organizations (CSO)/ implementation partners. •A gap in the data, as noted above, is the fact that there are no KIIS with, rightsholders/ beneficiaries or community level stakeholders. Secondary data collection sources are appropriate and well described - they consisted of longitudinal-observation design was used to review the data (documents for undertaking desk review) and sit reps on a monthly/ongoing basis to provide feedback and recommendations on each of the pillars.

	iii Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or, if not, provide reasons for this).	Partially	For KIIs with UNICEF staff, government officials and CSOs/ implementing partners there was a purposeful sampling strategy , developed in consultation with the CO, to ensure that the stakeholder list was appropriate and representative. The sampling strategy broken down by geographical district to ensure there was an appropriate geographical coverage. There is no reference to diverse perspectives , in capturing beneficiary or community level views: • <i>The evaluation did not collect data at the community level (with FLWs, teachers, health workers, SHGs, etc.) because reaching them virtually, especially during the emergency period, was not feasible for the evaluation agency (to conduct KIIs) as well as for the UNICEF Research and Evaluation team (to provide a list of potential stakeholders).</i> pp15 Given the fact that the evaluation team was operating under the valid premise of not adding to the burden of UNICEF and partners, this decision is understandable, in part. However, due to the importance of understanding the views of communities, and how/ if UNICEF was supporting them through this crisis, it may have been expected that greater effort was made to identify a small number of community stakeholders to speak with. For example, if UNICEF did not have contact details, for community stakeholders then CSOs/ implementing partners on the ground may have held them/ been able to provide phone numbers for people willing to be interviewed. In this way phone calls, WhatsApp or other commonly used messaging and voice calling services (available to many) could have utilised. A section outlining if this was explored as a means to capture diverse views would be useful to include.
	iv Clear and complete description of the methods of analysis.	Yes	The methods of analysis are clearly described.
	v Methodology allows for drawing causal connections between outputs and expected outcomes.	Not Rated	Not applicable as there was not a results framework from which to compare and contrast outputs and outcomes due to this being a Real Time evaluation.
	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	There is a clear description of the limitations and gaps in the evidence, and some mitigation strategies. However, not all limitations could be mitigated against, particularly the gaps in the data from community level.
	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).	No	There is mention of the UNEG standards, but no explicit or contextualised reference to them, or of how they apply to the evaluation team.
	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.	Partially	There is some mention of how the evaluation ensured avoidance of harm and protection of data from participants, but more specific and contextualised descriptions could have been provided.
	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 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.	Not Rated	Innovative practice was not used, or expected in the ToR.
	SECTION E: EVALUATION FINDINGS (weight 25%)		60%
	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 marshal sufficient evidence to systematically address most of the questions, but some are not addressed as during the evaluation process decisions were taken (in tandem in UNICEF CO) to not assess certain questions under different criteria, it states this is set out in section 2.2, but it is not clearly explained until the findings section. The questions which (presumably) lack evidence to assess primarily relate to meeting the needs of women and children, for example on Relevance the following sub question was not assessed: • <i>Provision of adequate health care for women, children, and vulnerable communities, including case management, provision of essential routine health and nutrition services</i> " (pp20).
	ii Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Not Rated	This was not an evaluation that analysed a theory of change but rather the process of response to COVID-19.

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.	Partially	<p>As this is a RTE output and outcome level data is not presented.</p> <p>There is reference to the evaluation matrix, which is relevant, and there is some use of triangulation of data. However, there is no triangulating from the community level, and only some from stakeholders. Evidence of triangulation could be improved:</p> <ul style="list-style-type: none"> • If more clarity were provided with regard to use of terms such as "some of the stakeholders" (pp19) or a "a few of the stakeholders" (pp21) -- vague descriptions do not allow an understanding of what type of stakeholder, how many stakeholders, or what geographic region they refer to. • Given that UNICEF staff have stated that they are not able to determine if supplies and services are reaching the desired populations, and they have to rely on local and state officials for evidence, if this evaluation had been able to reach community stakeholders they may have been able to fill a significant data gap for UNICEF in its response through triangulation of findings from different and situated perspectives. <p>Where the evidence is strong relates to:</p> <ul style="list-style-type: none"> • Providing evidence of support of UNICEF's alignment with, and effectiveness of, support to Government at state and district levels • Where information from UNICEF monthly situation reports is able to be triangulated by evidence from KIs with stakeholders and government officials.
	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	<p>Findings are supported by the evidence, where it is available, and it is noted throughout the findings section which questions can (or cannot) be answered.</p> <p>There is little means of comparison or indicators, which might have been expected, and this section could be improved if there was more comparative data, for example between different groups, or different geographical regions (states). Differences between the waves of COVID was mentioned, particularly in relation to effectiveness.</p> <p>Positive and negative findings are equally presented with the evaluation noting that there were some challenges in UNICEF's response and assumptions and estimates made in terms of supporting women and children:</p> <p>Negative: "The ... evaluation found that there was no formal mechanism to track or trace whether supplies and services were adequately reaching the target population." (pp24)</p> <p>Positive: "Digital mode of communication and training was adopted very well throughout the state by all stakeholders. As a result, the connectivity among different stakeholders increased." (pp44)</p>
	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).	Yes	<p>Causal factors were not expected to be easy to identify, given the nature of a RTE, and the lack of logic chain/ results framework. However, several causal factors were identified, such as:</p> <ul style="list-style-type: none"> • Overworked staff, within UNICEF and partners • Challenges with fundraising. finances to support activities • Lack of ability to track vulnerable people, and children. • Lack of ability to travel freely due to COVID 19
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	<p>The evaluation questions relating to the assessing the adequacy of the monitoring systems of UNICEF's intervention was removed as a part of the evaluation following the inception report, (pillar 4 of UNICEF's response - Data collection and social science research for public health decision making.</p> <p>The lack of effective monitoring systems for UNICEF's services and supplies is however, noted throughout the findings section, for example: "There was an absence of a comprehensive monitoring mechanism to track the progress of various activities and generate real-time learning to modify/alter the activities".pp38</p>
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).	Partially	<p>Conclusions are clearly formulated and reflect the purpose and objectives of the evaluation.</p> <p>They are not, however, particularly forward looking, and given this is a real time evaluation it would be expected they would have had significant elements and suggestions to improve the ongoing response.</p>

	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 conclusions are derived from the findings, and present the strengths and limitations. However, there is little insight and analysis is added beyond that which is already outlined the findings.
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.	Not Rated	Lessons are not presented, and the evaluation states they will be presented as a separate document, not included as part of this evaluation.
	ii	Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Not Rated	Lessons are not presented, and the evaluation states they will be presented as a separate document, not included as part of this evaluation.
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	Recommendations align with the evaluation purpose and are logically derived from findings. Overarching recommendations are provided, along with recommendations per pillar of UNICEF interventions, however, the pillar recommendations were made/ given to UNICEF at the completion of the evaluation of each pillar (during 2020-2021) so "Since pillar evaluations were undertaken at different time periods of the pandemic, the recommendations reflect the same. There may be some recommendations which may have been acted upon by UNICEF after the pillar-evaluation was concluded, and not be relevant in the current dynamic and ever-changing situation." (footnote, pp49). The recommendations could therefore be improved if there was clarity over which recommendations had already been implemented, or not, or a reason given as to why this information is not available. Recommendations should also be numbered.
	ii	Recommendations are useful and actionable for primary intended users and uses (relevant to the intervention); guidance is given for implementation, as appropriate.	Yes	Recommendations are useful and actionable for primary users (UNICEF), and guidance is given for implementation.
	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).	No	The process for developing the recommendations is not described.
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	Each recommendations is clustered by pillar, and has a note in brackets after it stating who is responsible for its implementation, however, this is not always clear, with some just stating the word "health" - is this the Ministry of Health, state level health, those in UNICEF or partners working on health specific interventions? Recommendations are prioritized, but as noted most are for immediate action, given this is a real time evaluation, and as some of the recommendations were made during 2020, there is no indication of those which have been completed, in part or in full. A table outlining this would be useful.
SECTION H:		REPORT STRUCTURE AND PRESENTATION (weight 5%)	79%	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.	Partially	Most of the expected information is included, however, the annexes are listed but an error has occurred and page numbers for the annexes are not provided, nor is it noted they are a separate document.
	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.	Partially	Annexes extend to 354 pages, and are very difficult to navigate given that they do not have a cover page with a contents and page numbers for each Annex. They do include key data one would expect, but the readability is poor due to use of large tables which span multiple (100s) pages. Additionally the ToRs appear to be a scanned copy, and as such are blurred and difficult to read.
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).	Yes	The report is easy to navigate and well formatted.
	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.	Yes	The evaluation follows the expected structure for a UNICEF evaluation report.
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.	Yes	The report is of a reasonable length.

	ii	Report is easy to understand (written in accessible way for intended audience) and generally free from grammar, spelling and punctuation errors.	Yes	The report is well written and generally free from grammar and spelling errors.
	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 good usage of figures and tables, however, while some figures and tables are named, some are not, for example the diagram on pp 10 is not labelled or numbered.
SECTION I:		EVALUATION PRINCIPLES (weight 10%)	19%	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.	No	There is no specific reference to rights based benchmarks and frameworks.
	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).	No	While it would not necessarily be expected that rights holders and duty bearers be as involved in a real time evaluation as other types of evaluation, it is expected that there would be some description of how rights holders were involved, even as part of data collection. Duty bearers were mentioned only as part of data collection processes.
	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).	Partially	Language used is inclusive, and challenges faced by marginalised groups is described. However, data is not disaggregated or analysed by marginalised groups, and "There was not sufficient data and information to make assessment of the gender outcome of the interventions, and comment on the gender transformative aspect of the interventions". pp46
	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	The evaluation does not explicitly mention child rights, but it did assess how well aspects of UNICEF's intervention met the needs of children in the key areas of activity (acknowledging data gaps). There is significant mention of the impact of COVID 19 on women, but a lack of disaggregated data to enable an assessment of how well the intervention met those needs. While there is mention of other excluded or marginalized groups, there is little meaningful engagement with these groups, or the data to enable it.
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)	1	
	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.	Partially integrated	GEEW was noted in the scope, and gender is noted as a cross cutting issue/criteria. However, this has not resulted in a gender appropriate scope of analysis, indeed many of the evaluating question relating to gender were not able to be answered.
	ii	A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Not at all integrated	The methodology was not gender responsive and there was little disaggregated secondary data. In addition the evaluation itself notes that as " the KIs were only with service providers and not with beneficiaries ... it was not possible to identify gender dimensions during the evaluation." (pp17)
	iii	The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Not at all integrated	There was some mention of GEEW as it related to service provision that was gendered in nature, for example services to pregnant women, and acknowledgement of the challenges faced by women during the COVID 19 pandemic. There was not however, a gendered analysis, as gendered data was not collected or integrated into the methodological 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>

iii 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?
- 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?
- c. Are unanticipated effects of the intervention on human rights and gender equality described?
- 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?