

GEROS Evaluation Quality Assurance Tool

Version: September 2021

Global Formative Evaluation of UNICEF Work on Disability Inclusion 2018-2022

REPORT RATING SUMMARY			
Overall Rating		54%	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 continue to use the evaluation with caution, but substantive improvements are possible.
●●●●●	Unsatisfactory (0% - 34.99%)	1	

REPORT DETAILS	
Title of the evaluation report	Global Formative Evaluation of UNICEF Work on Disability Inclusion 2018-2022
Report sequence number	Evaluation Office 65-2024-19890
Region	HQ
Year of report	2024
Office	Evaluation Office
Coverage (countries)	Global, but 14 case studies across all UNICEF regions
ToRs present	Yes
Date of review (dd/mmm/yyyy)	February 17, 2025
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	Yes
Gender equality (cross-cutting)	No
Humanitarian action (cross-cutting)	No
Evaluation object	Thematic area
Evaluation type	Summative and formative
Evaluation strategy	Mixed methods
Evaluation design (primary method used)	Theory-based
Evaluation level	Output & Outcome
Geographic scope	Multi-region/Global
Primary SDG(s) covered (number)	5, 10, 11

EOA 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.

This is a reasonable evaluation report, building on a good design and execution, though the analysis and findings are overly-superficial and could have been unpacked more to optimise the value of this important evaluation for UNICEF. Some of the key strengths are as follows:

- The report has a well-written and comprehensive context that effectively situates disability inclusion within global frameworks such as CRC, CRPD and internal policies such as UNDIS, DIPAS and UNICEF's strategic plans.
- A solid mix of methods incorporating primary/secondary, quantitative/qualitative (document reviews, key informant interviews, focus group discussions, and a confirmatory survey for triangulation) and a rights-based framing of the evaluation.
- A clear multi-stage sampling that included global, regional, and country-level perspectives as well as a good mix of duty-bearers and rights-holders.
- Findings are systematically organized according to the headline evaluation questions and all are answered.
- Recommendations are actionable and linked to institutional improvements.
- Good attempts to include gender-disaggregated data and intersectionality in disability analysis (despite challenges in data availability).

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.

The most significant issue with the report, which the rating reflects, is the somewhat limited granularity and depth of the analysis. While it is clear that considerable evidence would have been collected, many of the findings and analysis are not always sufficiently unpacked to provide UNICEF with maximum value in terms of how disability inclusion could be optimised. Another area that is not well represented is the voice of rights-holders - there is limited reference to the lived reality of disability from these key stakeholders, despite having been included in data collection. Some other areas which could have been improved are as follows:

- Information on the scope and focus of the evaluation is disjointed across multiple sections, with the scope section itself not including any details about it.
- The evaluation references the UNICEF ToC with respect to disability but does not systematically assess its effectiveness or reformulate it.
- Stakeholder mapping is limited – there is an absence of structured analysis of roles, responsibilities, and relationships among key actors.
- Findings are not organised under the evaluation criteria - they are presented under the headline evaluation questions only. The analytical elements of the evaluation matrix are not evident in the findings.
- Findings rely heavily on qualitative perceptions, with minimal cross-referencing between data sources to support a finding or to unpack causal factors that would be of significant value in a formative evaluation such as this.
- Limited assessment of UNICEF's M&E systems for tracking disability inclusion.
- There are several summary sections (5.2-4, Section 8) that do not add value to the report - they should have been integrated into the methods sections (e.g. a specific section on limitations) and the main findings sections.
- Recommendations are not ranked based on urgency/time-horizons or include any exploration of feasibility.
- While efforts to collect gender-disaggregated data took place, it is inconsistently applied, and a gender dimension is absent from the recommendations.

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 executive summary is six pages. This is one more than the TOR-mandated five pages, although only slightly, and given the wide-ranging nature of this evaluation, is not excessive. It is clear and concise and contains enough depth to serve as a standalone document useful for decision-making.
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 includes all of the necessary elements. The evaluation's purpose, objectives, and intended audience are briefly mentioned, as is the methodology. Key findings and conclusions are summarized, reflecting both achievements and challenges, and the recommendations are as per the main report.

	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	Insofar as the summary presents the headline findings, conclusions and recommendations exactly as they are in the main report, it does not introduce any additional information. However, some of the nuance (such as is available) under the findings and conclusions is lost as the headline elements are simply copied across - good practice summaries seek to synthesize the findings and conclusions more comprehensively.
SECTION B: BACKGROUND (weight 5%)			64%	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.	Partially	The object of the evaluation is somewhat clearly described across Section 2 of the report, although it is slightly disjointed in presentation, and thus not completely clear. Section 2.1 is to the-point in specifying the evaluation focus on assessing UNICEF's global work on disability inclusion in the context of two key policies related to disability: UNDIS and DIPAS. However, it references these two foundational policies without explaining the acronyms at what is their first usage in the main report. Even though they are used in the context of a quotation from the TORs, their nature should be explained using a footnote at this point. There is explanation of both under the following section (Logic Model, 2.2) but should be done earlier. The logic model (section 2.2) and the evaluation context (section 2.3) further explain the policy background, although it is somewhat duplicative between the two sections. Further, the sections lack a clear timeline of key milestones and policy shifts shaping UNICEF's disability inclusion strategy over the period. Implementation status is specifically covered under section 2.4, although briefly – the table of country examples could more usefully have been included here. There is no financial overview presented in these sections beyond noting the 2022 expenditure on disability only – this is covered in more depth under the findings, but might have been summarized in this section.
	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	As a high-level global evaluation, detailed itemization of the specifics of disability across national/sub-national levels is not appropriate, and the report does provide a high-level description of children with disabilities (CWDs) reached in section 2.4, with key stakeholders/duty-bearers outlined in the following section. This is very brief, however, with, for example, no disaggregation of numbers reached by gender, age or other intersectional vulnerability (ethnicity, humanitarian vs. development etc.).
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.	Partially	The report explains that the evaluation examines how UNICEF has promoted disability inclusion globally, regionally, and at the country level, aligning with key frameworks such as the Convention on the Rights of Persons with Disabilities (CRPD) and the United Nations Disability Inclusion Strategy (UNDIS). It also places disability inclusion across UNICEF's entire global reach of 190 countries and territories, although does not provide a structured breakdown of any systematic geographic focus of specific efforts beyond a table of examples from the selected country case studies (which, covering 1.5 pages, is on the extensive side and could have been condensed to <0.5 pages).
	ii	Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Yes	The SDGs and relevant targets are clearly noted under the relevant global policy instruments, with the report stating that "all SDGs are relevant to persons with disabilities", but itemising the specific SDGs (4, 8, 10, 11, 17) that are particularly relevant, i.e. have targets related to disability.
	iii	Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Yes	The introduction to the evaluation (section 1) and the following section and subsections (2.1, 2.2, 2.4) do discuss the status (i.e. prevalence of disability worldwide), with section 1 specifically noting many of the key issues and needs of CWDs worldwide. This is clear, concise and specifically relates to the evaluation.
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	Section 2.5 (pg. 25) provides a broad but not particularly specific identification of key stakeholders involved in UNICEF's disability inclusion work, including UN agencies, governments, bilateral and multilateral donors, OPDs (Organizations of Persons with Disabilities), private sector actors, academia, and CSOs. OPDs are mentioned as increasingly important partners. However, while the report lists stakeholders, it does not systematically map their roles and linkages. There is no clear stakeholder map or structured breakdown of how these entities interact across different levels (global, regional, national). Additionally, duty bearers (e.g., government agencies responsible for disability rights enforcement) are mentioned but not analysed in detail, or even examples of partnerships (as was done with respect to strategies and interventions in Table 1). A more detailed stakeholder mapping and an assessment of roles, accountabilities, and coordination mechanisms - both internal and external would have strengthened this section.
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Partially	As noted, section 2.5 provides a (useful) high-level summary of the broad typologies of stakeholders internal and external to UNICEF in relation to disability. The introductory sections provide some additional details on the contributions and roles of key stakeholders, including financial, technical, and programmatic roles, but this information is unevenly presented and lacks a structured analysis of stakeholder responsibilities. UNICEF's contributions are clearly defined in relation to policy advocacy, capacity strengthening, disability-inclusive programming, and data collection efforts and those of partners with in implementation, advisory roles and advocacy, but this is general and limited. The financial contribution of UNICEF to disability inclusion in 2022 (of \$285 million) is noted, but there is no systematic analysis of financial contributions until later in the report. In short, while the report identifies key stakeholders and provides some insight into their roles, it lacks further granularity and a more structured breakdown of contributions, roles/responsibilities etc. (even as an annex) would help the reader grasp the nature of the sector.
SECTION C: EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)			50%	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.	Partially	Inasmuch as the purpose is presented in the evaluation report, it is clear (it has both formative/learning and accountability elements). While the need for the evaluation at this point in time is justified by UNICEF's recognition of disability inclusion as a core strategic focus for the next decade and its commitment to aligning with the forthcoming strategic plan and global frameworks such as the newly-launched DIPAS, this is not made clear in the relevant section (3), which does not offer any explanation of the need for the evaluation at this point and time. Further, the purpose of the evaluation is articulated differently to that presented in both the TOR and the Inception Report - the substance is roughly similar, but, for example, the Inception Report purpose clearly mentions DIPAS but the evaluation report does not. The intended audiences and key users are well identified, with UNICEF's Disability Section, Programme Group, sectoral leads, and regional and country office focal points as primary users. Secondary users include external partners, OPDs, donors, and governments, with their specific use cases for the evaluation findings outlined in terms of programming and policy development.
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).	Partially	The objectives and scope of the evaluation are clearly stated and generally realistic, but there are issues with completeness. The evaluation objectives align well with the evaluation's stated purpose and are practical in scope, covering both development and humanitarian contexts. However, there is no discussion of the modifications made to the objectives from the original ToR during the inception phase – the final (presumably agreed-upon) set of objectives differ quite substantially from the original set. This should have been noted and the rationale/process explained.
	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).	No	The scope of the evaluation, which is well-defined in terms of geography, timeframe, and thematic coverage in the TORs, is not discussed explicitly at all in this section. While it can be inferred that it is defined as global, covering multiple programme areas and both development and humanitarian contexts, and the timeframe is clear from the introductory text, there is no explicit description in this section. Given that the section title explicitly mentions "scope", this is a clear omission that should have been caught at editing stage.
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 introductory sections (Sections 1, 2, and 3, Pages 16–29) provide a reasonably clear and structured description of UNICEF's work around disability, given that it is not a discrete programme but rather a broad area of work that is intended to be integrated across the organization. The evaluation report does include a graphical description of how disability is included in UNICEF's strategic plan (taken from UNICEF) which is analogous to a high-level theory of change (ToC) for disability inclusion (figure 2, pg. 6) and a more detailed discussion of this framework appears later in the annexes (although there are issues with this - disused in the next question).
	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	The underlying causal linkages to UNICEF disability inclusion are only partially discussed in the introductory sections. While the logic is represented at strategic level, there is no reference made to the evaluation ToC in Annex 4 therein, and no diagrammatic representation of the ToC is presented in either the Annex (which is a narrative approach) or the introductory sections. Further, the individual country case studies in Annex 9 all reference a ToC for the evaluation which has been based on the DIPAS ToC – but this is not included (or, if the ToC in Annex 4 is indeed the DIPAS ToC, this is not made explicit). Without a clear description of expected outputs, outcomes, and impact pathways in these early sections, it is difficult to determine which parts of the results chain are being tested by the evaluation. A clearer, better-structured articulation of the intervention's intended results, as well as an explicit linkage to the DIPAS ToC and annex 4, would have improved the clarity of the evaluation's analytical underpinnings.
	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	The ToC is noted explicitly as being subject to assessment in the evaluation, but its analysis is somewhat fragmented and lacks explicit reformulation by the evaluators. The evaluation acknowledges (in section 4) using a ToC "developed alongside DIPAS" and applies it to assess UNICEF's contributions to disability inclusion at the country level. However, the ToC and its development process are not clearly discussed in the report background sections (1–3) and the elaboration in Annex 4 is narrative only and not fully clear (the annex text notes that it "will be validated" vs. "was validated"). Further, the last sentence of section 4.3 (pg. 35) notes that the ToC was validated at the recommendations workshop, which comes at the end of the evaluation process, so it is difficult to see how a validated ToC could have been used for the evaluation design. Therefore, while the evaluation has developed or adapted a ToC and references it clearly, the process whereby it was validated and included in analysis is unclear, with no clear indication that the evaluators revised or reformulated it. There are clear inconsistencies in how the ToC was developed and used in the report which should have been resolved.
SECTION D:	EVALUATION DESIGN AND METHODOLOGY (weight 20%)	55%	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	Notwithstanding the absence of description of the revision of the EQs from those in the original TOR (discussed above), the evaluation questions presented in Section 3 are appropriate for meeting the objectives and purpose of the evaluation, as they do align with the (re)stated evaluation objectives and cover key aspects of UNICEF's disability inclusion efforts. The evaluation questions are explicitly aligned with OECD-DAC evaluation criteria standards and include relevance, coherence, effectiveness, and sustainability, although not organized under them (the criteria are presented separately). The discussion of the criteria, however, is appropriately to evaluate UNICEF's disability inclusion strategy, but questions could be more explicitly linked to each criterion to ensure a structured approach to analysis (as is done in the evaluation matrix in the annex).
	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 in Annex 6 breaks the evaluation questions down to 30 sub-questions that logically flow from each, as well as data collection methods, sources of data, and analysis techniques, also specifying data triangulation techniques including document analysis, thematic analysis, and descriptive statistics. The matrix also usefully includes a column for limitations related to each evaluation question which might affect data collection, which is an informative addition. The matrix has one key omission, however - there is no reference to the focus group discussions in the sources of data. There are two references to "Transcripts from key-informant interviews with children with disabilities" (EQs 3.2 and 4.2) but no more indication of the participation and/or contributions of rights-holders to data collection. The absence of rights-holders voices in the findings is perhaps a consequence of this (discussed below).
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	The evaluation design and methods in Section 4 (Pages 30–37) are generally relevant to the evaluation's purpose, objectives, and scope, albeit with some inconsistencies and presentation issues. The evaluation follows a theory-based, non-experimental, mixed-methods approach, incorporating contribution analysis, which is appropriate given the evaluation's goal of assessing UNICEF's influence on disability inclusion efforts The inclusion of qualitative and quantitative methods (document reviews, key informant interviews, focus groups, and a confirmatory survey) should have ensured good data triangulation. One inconsistent area is around gender, which the evaluators note was designed into the evaluation. However, the evaluation conclusions (section 8, pg. 69) highlight the inability of the evaluation to include any comprehensive gender analysis as it applies to disability, due to lack of data. Further, the section is not very well structured or organized - it is repetitive in places (e.g. the confirmatory survey is described several times across the subsections) and key aspects such as limitations are not presented in a single section, but across different subsections.
	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	The data sources used in the evaluation include both qualitative and quantitative methods, ensuring triangulation across multiple types of evidence and stakeholders at different levels. The report outlines three primary sources: (1) document review of internal and public UNICEF reports and data (2) key informant interviews (KIIs) and focus group discussions (FGDs) with UNICEF staff, external duty-bearers, and rights-holders (persons with disabilities) and (3) A confirmatory survey. These sources should provide a strong foundation for the evaluation. Section 4.5 is a useful discussion on the appropriateness of the methods, albeit brief - the table presented in the section would have been more useful if it had cross-referenced the tools as well as the data sources.
	iii	Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or, if not, provide reasons for this).	Partially	The sampling strategy is well explained in Section 4.3, outlining a multi-level sampling framework that considers UNICEF's organizational hierarchy (headquarters, regional, and country levels). The evaluation notes that 14 countries were selected for in-depth study, derived through a systematic process using mentions of disability inclusion in Country Office Annual Reports from 2018–2022. Four countries for in-person field work were selected on the basis of regional balance and recommendations from Regional Offices (i.e. purposive/convenience sampling). While the sample process is well described here, the selection process for countries (essentially, those that were actively engaged in disability inclusion work were prioritized) introduces a significant positive bias. Including countries that performed less well (i.e. struggling to initiate work) could have provided useful comparative insights. The limitation of not including countries with weaker disability inclusion efforts is noted (on pg. 34), but not addressed. Further, the survey tool, which was extended to 150 countries (and could therefore have mitigated the selection bias for the other tools) was responded to on a voluntary basis – only 48 countries completed this survey – again, a likely source of bias. On an individual stakeholder basis, while diverse perspectives were included through interviews and FGDs with UNICEF staff, external duty-bearers and rights-holders, the sampling process for these is not described. A more explicit discussion on how diverse perspectives were ensured or why certain perspectives were excluded would have strengthened the credibility of the evaluation findings.
	iv	Clear and complete description of the methods of analysis.	Partially	There is a brief description of the data analysis process provided in Section 4.1. It covers the basics of the data analysis process, noting some of the broad principles underlying the data analysis (triangulation, inductive and deductive processes) and the general analytical approach (descriptive analysis of quantitative data) and use of NVivo and MS Excel software, but there is little else. In particular, there is no discussion of any data compilation, cleaning, coding etc. that would explain how data collected was linked to the analytical framework/evaluation matrix in a systematic manner. Thus the section is minimally adequate, but would have benefitted from more comprehensive detail. There is more detail regarding the contribution analysis process presented later in the section and in Annex 5, which is positive, but this should have been included in Section 4.1 rather than 4.3, which covers sampling.

	v	Methodology allows for drawing causal connections between outputs and expected outcomes.	Yes	The description of the evaluation methodology, variety of tools and stakeholders, plus the contribution analysis approach (described under section 4.3, supported by additional detail in Annex 5) would suggest the capacity to draw of causal connections between outputs and expected outcomes, albeit with some minor limitations. The contribution analysis approach is appropriate for understanding UNICEF's role in disability inclusion in that it examines whether UNICEF's policies and implementation strategies have influenced programme results and attempts to strip back rival explanations and therefore establish direct causal linkages between specific outputs and their intended outcomes. The evaluation also acknowledges external influencing factors (e.g., government commitments, funding availability, and social attitudes toward disability), identification of which will help to more clearly attribute observed changes to UNICEF interventions.
	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	Although section 4.1 does specifically note limitations in the section heading, the report does not include a specific limitations section, which is against generally-held good practice. Instead, limitations are noted across various sections of the report, including mentions in the methods (section 4.1), the findings (section 5.3) and the conclusions. This is somewhat disjointed. Where included, the report does discuss the limitations and constraints faced by the evaluation, but it is more of an acknowledgement of inherent weaknesses in the data sources (for example the potential lack of nuance in published reports, the situational nature of stakeholder perspectives, and the limited representativeness of interviews) rather than a systematic approach that notes the limitations, efforts to mitigate them, and the ultimate outcome on the evaluation itself. Overall, while the limitations are acknowledged and some mitigation strategies are presented, there is a lack of explicit discussion on whether remaining gaps in evidence significantly affect the reliability of conclusions.
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).	Partially	The evaluation report explicitly references the UNEG Ethical Guidelines for Evaluation, and lists the key principles (utility, credibility, independence, impartiality, ethics, transparency, human rights, gender equality, national evaluation capacities, and professionalism) in section 4.6. Additionally, it confirms adherence to the UNICEF-adapted UNEG Evaluation Standards and UNICEF's own standards for disability-inclusive evaluation. However, while the report asserts compliance with UNEG obligations and principles, it does not provide in-depth discussion or contextualized examples of how the principles were managed in practice, per the criterion.
	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	While section 4.6 provides general affirmation of adherence to various ethical policies and safeguards (which are linked in footnotes), there is limited discussion of how specific risks were identified or mitigated. The section does highlight that informed consent was ensured in accessible formats, and participant privacy, confidentiality, and autonomy were respected. It also indicates that evaluators took responsibility for ethical considerations throughout the evaluation cycle. However, while the TOR confirmed that no independent ethical review was required, there is no explicit mention of the UNICEF procedures for Ethical Research Involving Children, which would have been particularly relevant given that children with disabilities were included in the evaluation. Further, given that children were intended to be interviewed on their own (per guidance given in data collection Tool L), independent ethical review of the tools <u>should</u> have been sought, <u>contrary</u> to UNICEF direction in this regard. Additionally, while ensuring representation of marginalized groups is emphasized, there is no discussion on whether additional protective measures were taken for particularly vulnerable participants who may have required tailored safeguarding protocols. Finally, the data collection tools are not annexed to the report, so determining their compliance with ethical processes necessitated a review of the Inception Report (to which they are annexed) - in this there is good adherence to consent procedures.
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.	No	There is no indication of any innovative practices in the evaluation design or implementation.
SECTION E: EVALUATION FINDINGS (weight 25%)			42%	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	Firstly, the evaluation findings are presented solely against the evaluation questions (of which there are eight). The sub questions present in the evaluation matrix are not evident in the analysis, and while the DAC criteria that relate to each EQ are noted, the findings are not organised according to the criteria (there is a one-page subsection (5.2) that specifically discusses alignment with the specified criteria, but this is more of a conclusions statement, as there is no analysis or evidence presented here). Secondly, the evaluation criteria are not fully consistent with what is in the evaluation matrix. For example, EQ1 is aligned with the Relevance criterion in the main report, but is noted as being "descriptive" (i.e. no DAC criterion alignment) in the Matrix. There is something similar with EQ2 and EQ7. Thirdly, the levels of evidence presented against each evaluation question vary across questions, with levels somewhat thin (the findings section is 21 pages, which, while a crude measure, is not a considerable quantity for an evaluation of this nature). That said, the findings clearly correspond to each evaluation question and draw on multiple data sources such as interviews, survey (and secondary data) analyses and country-level examples. The findings articulate specific evidence (for instance, citing instances of policy advocacy, system strengthening and identification of children with disabilities) that connects directly to the EQs and criteria. The findings do acknowledge constraints such as data gaps, but consistently reference triangulated information – including qualitative inputs from rights-holders, thus providing a credible basis to address the scope of the evaluation.

	ii	Explicit use of the intervention's results framework/ToC in the formulation of the findings.	No	There is no explicit reference to the theory of change developed during the evaluation in the findings section. While the findings align implicitly with the evaluation's objectives and thematic questions, the text does not explicitly connect the analysis to the stated ToC. The findings section uses the evaluation questions and criteria to structure its presentation of evidence, but does not specify how each finding links back to any intended results framework.
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	The report presents reasonable forms of qualitative and quantitative data across the range of findings. For example, under EQ7 (section 5.1.7), there is a synthesis of qualitative inputs from country offices and quantitative financial data. It is clear that findings are derived from multiple country case studies (with many country-specific examples cited), showing how broader systemic changes were influenced by UNICEF interventions, which are corroborated by survey analyses and secondary data/documents. However, while the report does make use of multiple forms of data for findings, actual triangulation of findings - i.e. using multiple sources to deepen analysis explore causality, is quite limited. In many cases, the text discusses stakeholder interviews and survey insights in general terms but does not lay out side-by-side comparisons or clearly show matching trends that correspond across these different sources - while the report alludes to a diversity of evidence, the analysis itself does not offer a strong, explicit illustration of how each data type supports or strengthens the other. For example, under EQ4 (5.1.4), survey results on the use of disability data are simply presented with quite limited any unpacking of what this means. This is discussed further under 13iii.
	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 findings are generally supported by evidence, but the degree of clarity and substantiation varies. Many findings reference qualitative sources, (primarily key informant interviews, with very limited reference to rights-holder views via focus group discussions), but they are often presented in a generalized manner without direct attribution. Quantitative data is used (mainly the survey results, with some secondary quantitative data analysis), but it is not consistently used to benchmark findings or substantively triangulate or unpack findings. There is limited explicit use of formal benchmarks or comparative performance measures, making it difficult to assess the extent of progress relative to predefined goals. The evaluation findings are presented under the eight evaluation questions only, with no more specific itemization under sub questions (and, in the case of EQ7 (5.1.7), the specific parameters of "(i) partnerships, (ii) financing and resource mobilisation, (iii) leadership accountability, (iv) coordination (v) cross-sectional)" noted in the evaluation matrix are not referenced systematically in the analysis). While some findings acknowledge both successes and gaps (e.g., on policy uptake or implementation barriers), there is little structured comparison against expected results or global best practices. The evidence presented suggests validity, but it is somewhat superficial, with the findings relying more on thematic synthesis than on structured, systematic and granular performance assessment.
	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	The findings do identify some causal factors underlying the analysis, but the depth of analysis varies. Organizational and managerial factors, such as the role of internal champions (EQ7), the "invisibilisation" of children and their lack of participation (EQ4) and resource availability (EQ7), are frequently mentioned as enablers or constraints to disability inclusion efforts. However, while these are acknowledged as influencing factors, their specific mechanisms or interactions with other elements (e.g., policy frameworks, external partnerships) are not consistently elaborated, with many of the findings somewhat thinly argued as a result. Contextual factors, including country-specific challenges, policy environments, and socio-cultural barriers, are referenced in some findings but are not systematically explored in relation to their impact. The report acknowledges implementation gaps, such as limited technical guidance on the "Twin Track" approach, but does not always clarify whether these gaps stem from internal strategic limitations, external constraints, or a combination of both. Finally, some of the survey findings are simply presented as they are, with little or no unpacking of the underlying rationale - for example, under EQ3 (pg. 44) the survey results on perceived effectiveness (or lack thereof) of UNICEF policy advocacy work by staff are presented, but the reasons for the results (i.e. triangulated with richer interview data) are not explored. With respect to the logical chain, some findings describe how UNICEF's interventions—such as policy advocacy, technical assistance, and capacity-building efforts—contribute to disability inclusion outcomes. However, the report does not consistently map these efforts to a clear results chain, meaning that the connection between specific activities and their eventual impact is often implied rather than explicitly demonstrated.
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	There is some, but limited assessment of the adequacy of UNICEF's efforts to monitor disability programming, notably under EQ5, which highlights the shortcomings of disability disaggregated data availability globally. This section does reference data gaps and challenges in tracking disability-inclusive programming, but does not discuss any results/performance framework or the presence/absence of a monitoring system in this regard - which is a missed opportunity. Some findings mention the use of disability-disaggregated data and reporting mechanisms such as the development of the Washington Group Child Functioning Module, this is an external data initiative rather than a structured analysis of how well UNICEF's M&E systems (such as the UNICEF RAM) integrates disability data to support decision-making. Key components such as the alignment of indicators with strategic objectives, the reliability of data collection tools, or the extent to which monitoring data informs adaptive programming are not explicitly addressed.
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>Firstly, the report presents a number of closing analyses prior to the conclusions section (sections 5.3 and 5.4) which are atypical of evaluation reports and are largely (as are the conclusions) a structured itemization or summary of findings. Section 5.4, in particular, notes its inclusion as "United Nations Evaluation Group (UNEG) recommends a section that outlines reasons for "recommendations and failures" – this appears to be a misinterpretation of UNEG guidelines, which specify that findings should include "reasons for accomplishments and failures", not that a specific section be included.</p> <p>Within the specifically relevant section (6) the conclusions themselves are formulated with clear reference to the evaluation's core themes, emphasizing the strengths, opportunities, and bottlenecks encountered in UNICEF's disability inclusion efforts. They are somewhat forward-looking in their orientation towards UNICEF's future disability programming in that they recognize major gaps and challenges and mention pathways for improvement (e.g., clarifying the "Twin Track" approach and strengthening internal capacity). This provides a link to the recommendations that should specify operational or strategic adjustments needed within UNICEF and how UNICEF should adapt its programming and systems in the future.</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	<p>The conclusions are broadly aligned with the findings but do not consistently extend the analysis to provide deeper insights beyond them. They effectively synthesize key themes, such as UNICEF's rights-based approach to disability inclusion, the role of internal champions, and challenges related to the "Twin Track" strategy. However, they primarily restate findings rather than offering new analytical perspectives that critically examine underlying causes or systemic constraints. For example, while gaps in institutionalization and data collection are mentioned, there is little further discussion on why these gaps persist or what structural barriers within UNICEF or its operating environment contribute to them. As a result, while the conclusions accurately reflect the evaluation's findings, they do not provide an enhanced level of insight – this is somewhat reflective of the findings themselves, which are in places superficial and lacking deeper analysis of underlying causes.</p>
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.	Partially	<p>No lessons-learned section was requested in the TOR - instead, lessons were ostensibly integrated across the evaluation objectives and questions, and in the overall analytical design of the evaluation. This said, the objective of documenting lessons about disability - an original component in the TORs (objective g) - did not form part of the four restated objectives in the Inception Report nor of the evaluation questions (though is covered by multiple sub-questions in the evaluation matrix, although the report does not refer to these in its findings). The lessons included do have wider applicability and relevance to disability-inclusive programming.</p>
	ii	Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Partially	<p>There are various lessons-learned/good practices scattered throughout the report, which are positive, but would have been better compiled in one section.</p>
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	<p>The four recommendations are relatively clearly formulated and do align with the overall intent and purpose of the evaluation insofar as they address UNICEF's capacity to mainstream disability and inclusion and highlight some key deficits/opportunities in this regard. However, they only partially address all the critical issues identified in the conclusions (and are not explicitly aligned with individual conclusions). For example, Conclusion 7 acknowledges that while UNICEF has made progress in integrating assistive technologies, their impact is constrained by inconsistent implementation and structural barriers. Yet, the recommendations focus primarily on improving procurement and distribution mechanisms without detailing how UNICEF should ensure the effective utilization of these tools within enabling environments. The environment that enables assistive technology - noted clearly in the conclusions - is only alluded to briefly in recommendation 1.</p> <p>Similarly, while Conclusion 3 highlights gaps in technical guidance and disability-disaggregated data, the corresponding recommendation on the "Twin Track" approach (#2) and on global advocacy (#3) focuses on general efforts rather than addressing the root causes of these gaps - disaggregated disability data is already a commitment by UNICEF - the evaluation should have identified, and recommendation(s) should focus more on addressing, the underlying causes why this data is not comprehensively collected.</p>
	ii	Recommendations are useful and actionable for primary intended users and uses (relevant to the intervention); guidance is given for implementation, as appropriate.	Partially	<p>The recommendations are broadly useful and actionable for the primary intended users. They provide clear, if somewhat generic (e.g. Recommendation 1) direction on institutionalizing disability inclusion, improving technical guidance, enhancing data collection, and strengthening advocacy efforts.</p> <p>However, the level of specificity varies, and some recommendations lack detailed implementation guidance. For example, recommendations related to institutionalization (e.g., establishing accountability mechanisms and budget lines) are concrete and clearly targeted at different levels of the organization. However, other recommendations, such as strengthening technical guidance for the "Twin Track" approach, are more general and do not provide sufficient details on how this guidance should be structured, disseminated, or integrated into existing UNICEF frameworks. Further, the recommendations also assume that UNICEF has the internal capacity and resources to execute these actions immediately, without explicitly addressing potential barriers to their adoption at scale</p>
	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	<p>The evaluators (in Section 4.3, which describes the Sampling process, so is somewhat misplaced) describes a process for developing recommendations, involving a co-creation workshop with the evaluation team, the Evaluation Office, and Programme Group staff. This workshop aimed to validate the draft theory of change and finalize the recommendations. The description of the recommendation development process does not include external duty-bearers or rights holders and there is no stated explanation for why rights holders were not involved in this phase, if that was the case.</p>

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 evaluators clearly identify responsible groups or duty-bearers for most actions. Each recommendation specifies the primary entities responsible for implementation, such as the Disability Section, Programme Group, regional and country offices, and other divisions like the Division of Global Communication and Advocacy and the Supply Division. Regarding prioritization, all recommendations are given an "Immediate" timeframe, but there is no further classification based on importance. This limits the ability of decision-makers to determine which actions should be prioritized. Additionally, while the recommendations specify adaptations for regional and country-level implementation, they do not explicitly outline sequencing or dependencies between recommendations, which could enhance their usability for strategic planning.
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.	Partially	Most of the required elements are present, but the evaluators are not apparent from the opening pages. The team members are named in the preface text only. The table of contents/list of tables/figures is also present, but has errors - noted below.
	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	Most of the specified annexes are included, excepting the data collection tools (which were included in the Inception Report but not in this report - their extent may have precluded inclusion here). The annexes also include additional methodological details, additional analyses and the individual country case studies.
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	The formatting is mostly acceptable, with no issues with the numbering, the titling etc. However, the box/table/figure numbering system in the list of tables/figures is incorrect - Box 2 is not bookmarked and Box 3 in the list is actually Box 2 in the text and vice versa. Similarly for the tables (inconsistent numbering of tables and page numbers not matching up, and also for the figures, which also have a "bookmark not defined" error.
	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	The broad structure follows UNICEF guidelines, but there are additional sections that are not part of generally-accepted evaluation report, notably additional quasi-conclusion or summative sections 5.2-5.4 and section 8. These do not add to the report and should have been considered for removal.
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 TOR specifies 60 pages for the evaluation report (excluding annexes and the executive summary). This report is 54 pages, so meets the standard comfortably - indeed, it could be further shortened by removing some of the subsections that are of questionable added-value (see previous point).
	ii Report is easy to understand (written in accessible way for intended audience) and generally free from grammar, spelling and punctuation errors.	Yes	Overall, the report is written in an accessible manner, with clear language suitable for its intended audiences. The structure follows a logical flow, making it relatively easy to navigate. There are some minor errors remaining, but no significant spelling or punctuation errors that undermine readability.
	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	The report includes various tables and figures to support the presentation of key information. Tables summarize data on sectoral interventions, country-level programming, funding for disability inclusion, and policy advocacy outcomes. Additionally, figures/charts illustrate elements related to quantitative data (e.g. the survey results and funding trends). While these visual aids are generally clear and labelled, some figures are referenced generically in the text without detailed discussion or integration into the narrative. Additionally, at least one figure contains a labelling issue ("Error! Bookmark not defined."), which suggests that some formatting inconsistencies may affect readability and some figures (e.g. Figure 5) have the caption on one page and the chart on another. Finally, at least one of the figures (e.g. Figure 2) is too low resolution to read properly. The report could improve the use of visual aids by ensuring that they are clearly presented, formatted, explicitly referenced and analysed in the text to enhance comprehension and reinforce key points.
SECTION I:	EVALUATION PRINCIPLES (weight 10%)	79%	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	The report explicitly references and incorporates rights-based frameworks from the outset, most importantly the CRC and the CRPD, and CEDAW, all include articles specific to children and women with disabilities, particularly in the framing of UNICEF's disability inclusion approach. These frameworks are mentioned in relation to UNICEF's commitment to disability inclusion as a cross-cutting issue, aligning with broader international human rights obligations. They are reflected in the evaluation design via the evaluation matrix.

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	An Evaluation Reference Group, consisting of UNICEF staff and external partners, was convened at the inception phase to contribute to the design and conceptualization of the evaluation. The design of the evaluation highlighted consultations with staff from headquarters, regional offices, and country offices, as well as with global partners identified by the Disability Section. Additionally, field-based data collection included interviews with UNICEF personnel, governmental and NGO stakeholders, community representatives, parents of children with disabilities, and children with disabilities themselves. This is clear and appropriate (although the voices of rights-holders is not particularly apparent in the findings).
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	The report generally uses inclusive and empowering language, with a strong rights-based framing that aligns with UNICEF's commitment to disability inclusion. It avoids overt biases related to gender, age, culture, or religion, and maintains a neutral, professional tone that is appropriate for its intended audience. However, while the report applies a rights-based approach conceptually, it does not consistently use terms like "rights holders" and "duty bearers" throughout. Instead, for the former, the term "beneficiaries" is more common, and for the latter, it refers more generally to UNICEF, governments, and implementing partners without systematically terming them as duty-bearers. Additionally, while data disaggregation by gender is present in some sections, the evaluators acknowledge a gap in intersectional analysis, particularly regarding the differential experiences of girls and boys with disabilities. The evaluators note that data limitations prevented a more nuanced gender-disability analysis.
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.	Yes	The evaluation (which is focused on disability inclusion, so clearly meets this sub-criterion) explicitly assesses the extent to which UNICEF's disability inclusion efforts align with child rights and the Leave No One Behind (LNOB) principle. This is reflected in both the main analysis, and can be seen in the individual country analyses in the annexes. It clearly acknowledges the importance of reaching marginalized children with disabilities, particularly in education and early childhood development, and is grounded in various rights-based frameworks (noted above).
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)		6	
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	The evaluation integrates GEEW within its scope of analysis, using a human rights-based perspective and a gender lens to assess the implementation of disability-inclusive programming. The evaluation criteria and questions were designed to capture gender-related data, and the methodology explicitly includes gender equality as a principle for data collection and analysis
ii	A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Satisfactorily integrated	While gender considerations were incorporated in the evaluation approach and tools (e.g. interview/FGD questions probed for differential experiences among boys/girls/men/women, it was not the primary focus of the evaluation. Further, the report acknowledges that intersectional analysis of disability and gender was constrained by data limitations. Although representation of males and females in case studies was equal, the available data did not allow for a thorough assessment of how disability interventions differentially impact girls and women. Furthermore, gender-disaggregated data is inconsistently presented across findings, which limits the ability to fully assess UNICEF's performance in addressing gender equity within disability programming.
iii	The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Partially integrated	The evaluation findings, conclusions, and recommendations incorporate some aspects of gender analysis but not fully consistently. The findings acknowledge gender as a factor in disability inclusion but lack a systematic assessment of gender-disaggregated results across UNICEF's interventions (discussed above). While the conclusions highlight the need for intersectional approaches, particularly concerning gender and disability, the evaluation itself does not comprehensively explore how UNICEF's disability programming impacts boys and girls differently. The recommendations mention improving disability-disaggregated data but do not explicitly call for stronger gender analysis within UNICEF's monitoring frameworks. Additionally, while the recommendations emphasize reducing stigma and invisibility for children with disabilities, they do not note the specific barriers faced by girls/women with disabilities and, indeed, gender is not referenced in the recommendations.

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>