

1.2. Methodology

Criteria of Quality

The criteria of evaluation quality applied in this study (see Box 1 and Appendix 1), and the questionnaire by which information was gathered from UNICEF country offices (Appendix 2), were identified from the published evaluation standards of professional evaluation associations; from literature³, especially previous reviews of evaluation quality by UNICEF⁴, USAID⁵ and CIDA⁶, and from discussion among the professional evaluators in UNICEF and within the consulting team.

Sample of Reports

The criteria of quality were applied to a 50% random sample of UNICEF evaluation reports submitted to New York Headquarters by the country offices in the years 2000 and 2001 (approximately 97 reports, of which 75 proved to be evaluations, strictly defined). The sample was drawn from a list of evaluations that had first been reviewed to sort the true evaluations from other types of research.³¹ As well, the assessment team examined the quality of 31 Terms of Reference for our sample of evaluations — about one third — all that were available at New York headquarters for this set of studies.

Assessment Procedures

Five consultants — all expert evaluators — assessed the quality of the UNICEF evaluations. (See Appendix 6.) . The team leader did the first five assessments to benchmark the standards of quality. Each evaluator reviewed a benchmark assessment, and completed an initial assessment of one evaluation that was discussed with the team leader in detail. The evaluators then proceeded to complete the set of assessments, discussing issues as they arose.

Reports were randomly assigned to assessors, except in so far as the language of the report was a constraint. Only one evaluator was competent to assess evaluation reports in Spanish, and one in French.

Assessors were engaged primarily for their expertise as evaluators and evaluation methodologists, but, in fact, the team did have wide experience in the main areas of UNICEF's work as well.

Survey of Country Offices

A questionnaire was sent to staff in the country offices — either the UNICEF country representative or the evaluation focal point in the office. The questionnaire was developed in conjunction with UNICEF headquarters staff, and tested with three country office focal points before it was finalised. (See Appendix 2.)

A 50% random sample of country offices received the survey. The reason for a sample was to minimise the response burden. By careful follow-up, the study team achieved virtually a 100% response. Only one country office declined to respond.³²

³¹ Research was considered an evaluation only if it examined an actual or potential investment or other intervention by UNICEF. By this measure, some research reports were misclassified as evaluations and were put aside.

³² One office declined to respond on the grounds that it had done no evaluations during 2000 and 2001, no staff currently in the office had been there during that time, and no current staff had evaluation experience.

BOX 1: CRITERIA OF EVALUATION REPORT QUALITY

General format of the evaluation report:

Was there a cover page and title? Was the report dated, and the author(s) named? Was there a table of contents? An Executive Summary? Were Terms of Reference attached?

Criteria of Quality

1. The project, programme, or initiative to be evaluated should be clearly described, including the logic of its links between activities, outputs, and outcomes.
2. The degree to which the project, programme, or initiative might be replicable in other contexts should be described.
3. The incremental contributions of UNICEF and other stakeholders, in cash or in kind, should be stated.
4. The context should be analysed in sufficient detail to identify external factors contributing to success or failure, and this analysis should be used to support informed judgements about what results may reasonably be attributed to UNICEF's intervention, and what to other factors.
5. The objectives of the evaluation, and the questions to be answered, should be stated fully and clearly.
6. The objectives of the evaluation, and the questions to be answered, should reflect UNICEF's mission and approach to programming, including protection of children's rights, promotion of their welfare, and gender equality.
7. The evaluation should use standard UNICEF evaluation criteria: relevance, efficiency, effectiveness, impact and sustainability.
8. The evaluation report adequately addresses the issue of sustainability.
9. The evaluation report adequately addresses the issue of relevance.
10. For evaluations of humanitarian response, in particular at the sector or whole country level, the evaluation should address coverage, coordination, coherence and protection.
11. The methodology of the evaluation should be practical and appropriate to the questions posed, while providing a complete and fair assessment.
12. The evaluation design should be ethical and include ethical safeguards where appropriate, including protection of the dignity, rights, and welfare of human subjects, particularly children, and respect for the values of the beneficiary community.
13. The evaluators should make a statement about how objectivity and independence were ensured.
14. The constraints of the evaluation, and the perspective from which the intervention is evaluated, should be clear so the reader can assess the validity of the evaluators' judgements.
15. The evaluation report should describe the kinds of UNICEF and other key stakeholders' follow-up action expected in regard to lessons and recommendations.
16. The information gathered, qualitative and quantitative, should, in aggregate, be adequate to answer the evaluation questions.
17. Sampling is used when appropriate and is designed to produce representative data.
18. The data collection instruments should be able to provide the measurements needed.
19. Information should be gathered from eligible persons not reached in addition to those reached.
20. The report should distinguish clearly between implementation factors [inputs, process milestones, activities] and results [outputs and outcomes/impacts].
21. Inputs [costs/resources] should be measured.
22. Outputs [what is actually produced by the inputs] should be measured.
23. Outcomes/impacts [what the intervention seeks to achieve] should be measured or an appropriate rationale given why not.
24. Costs and results should be directly compared to as full an extent as possible.
25. Accountability for results [who is accountable for what, and their performance] should be appropriately analysed and reported.
26. Lessons learned should be generalised beyond the immediate intervention being evaluated to indicate what wider relevance to UNICEF there might be.
27. Recommendations should be firmly based on evidence and analysis.
28. Priorities for action should be made clear.
29. The evaluation report should be complete, with all of the necessary elements of a good report present.
30. The evaluation report should be clear, transparent and easily accessible to the reader. It should be concise, well organised and logical. The text should be clearly written, supported by tables, figures and descriptive headings, and led by an executive summary.