Concerning these criteria, the evaluation report is: | Unacceptable | Poor | Good | Very Good | Excellent |
--- | --- | --- | --- | --- | --- |
1. **Meeting needs**: Does the evaluation adequately address the information needs of the commissioning body and fit the terms of reference? |  |  |  | x |  |
2. **Relevant scope**: Is the rationale of the policy examined and its set of outputs, results and outcomes/impacts examined fully, including both intended and unexpected policy interactions and consequences? |  |  |  |  | x |
3. **Defensible design**: Is the evaluation design appropriate and adequate to ensure that the full set of findings, along with methodological limitations, is made accessible for answering the main evaluation questions? |  |  |  |  | x |
4. **Reliable data**: To what extent are the primary and secondary data selected adequate. Are they sufficiently reliable for their intended use? |  |  |  |  | x |
5. **Sound analysis**: Is the quantitative and qualitative information appropriately and systematically analysed according to the state of the art so that evaluation questions are answered in a valid way? |  |  |  |  | x |
6. **Credible findings**: Do findings follow logically from, and are they justified by, the data analysis and interpretations based on carefully described assumptions and rationale? |  |  |  |  | x |
7. **Validity of the conclusions**: Does the report provide clear conclusions? Are conclusions based on credible findings? |  |  |  |  | x |
8. **Usefulness of the recommendations**: Are recommendations fair, unbiased by personal or stakeholders’ views, and sufficiently detailed to be operationally applicable? |  |  |  |  | x |
9. **Clearly reported**: Does the report clearly describe the policy being evaluated, including its context and purpose, together with the procedures and findings of the evaluation, so that information provided can easily be understood? |  |  |  |  | x |

Taking into account the contextual constraints on the evaluation, the overall quality rating of the report is considered |  |  |  |  | x |
Taking into account the contextual constraints on the evaluation, the overall quality rating of the report is considered: VERY GOOD

This assessment is supported by the following elements:
- Meeting needs;
- The soundness of the methodological tools used and the reliability of the data collected;
- The clarity and usefulness of conclusions and recommendations;
- The extensive, greatly detailed and in depth analysis;
- The width of the topic tackled, and the great effort made by the consultants to respond to the requests made in the ToRs

1. Meeting needs: VERY GOOD

The evaluation addresses all the information needs of the commissioning body and fits the Terms of References. The evaluation questions are answered in a satisfactory way and provide the reader with a very complete overview of the EC support to basic and secondary education. The extensive analysis conducted in the country notes goes beyond the request of the ToRs. The consultants, indeed, presented extremely detailed findings divided by EQ for each of the 6 Countries concerned and provided us with an exhaustive country profile (annexes 6-11). Moreover, in the annexes, we can find a thorough analysis of the international policy framework in the education sector and how the EC supports it (annexe 18), a note on international indicators (annexe 22) and an overview of international surveys related to learning achievements (annexe 25).

2. Relevant scope: VERY GOOD

The report addresses the entire scope defined in the ToRs in its temporal, geographic and regulatory dimensions. A limited but representative number of countries for desk study, field mission and focus groups were chosen as sample. The selection extensively covers the EC contribution to the education sector in terms of geographical and sub-sectoral distribution, participation to the FTI scheme, MDG contract, mix of modalities and the country specific context (fragile county, LRRD etc). Other donor's policies and support are looked at when dealing with multi donor initiatives and trust funds (Education For All, Fast Track Initiative etc). Unintended effects are mentioned, in particular when presenting conclusions (for example the impact of the support to basic education on secondary education).

During the evaluation process, as an addendum of the ToRs, evaluators were also asked to collect and examine additional data through focus groups and through the analysis of the European Court of Auditors report. These analyses were included in the report and contributed to further deepen and broaden the scope of the evaluation.

3. Defensible design: VERY GOOD

The evaluation methodology is very clearly and thoroughly explained both in the main report and in the annexes (Vol. 2d). Very detailed information is also provided on the inventory of
the EC support and on the approach that the consultants developed to undertake it (both for direct and for indirect support, see Vol.2b). Limitations inherent to the evaluation method, scope, processes and tools have been clearly specified and analysed, for example challenges and limits related to the inventory or to the country visits (p. 14 of the main report). The choices made and approaches developed to deal with limits and challenges have been thoroughly explained (p. 14 of the main report for limits related to the inventory, p. 27 and 37 of Vol. 2b for limits related to the CRIS database and inventory for GBS).

The team shows a great knowledge of and capacity to use the EC evaluation methodology and tools. The evaluation is structured in different phases and makes good use of evaluation questions, judgement criteria (JCs) and indicators. The tools employed are: document reviews, interviews, analysis of previous existing evaluations/monitoring reports, country level analysis for 23 desk study countries and focus groups.

The structure chosen for interviews and focus groups is presented in details in the annexes.

4. Reliable data: VERY GOOD

All data used for the evaluation are reliable and were extrapolated from trustworthy sources as: EC strategy and programming documents, documents by other international organizations, previous evaluations, result-oriented monitoring reports (ROMs), interviews etc.

The types of documents and data used are listed very clearly, on the basis of topics and geographical coverage. Although it could have been useful to have more punctual references for the information collected at JC or indicator levels, we consider that this has not impaired the reliability and the validity of data and of their use. The limits pertaining to available information and collection tools are explicit (p.14 Main report and p. 27, 37, 135 Vol. 2b). Information has been cross-checked and verified through different information sources and employing several tools that are independent of one another (surveys, case study, interviews, briefings with EU Delegations (EUDs) and stakeholders in the field, document and data analyses). The data collected through the EUDs survey and its contribution to the analysis cold have been highlighted a bit more.

5. Sound data analysis: GOOD

Volume 2a shows in details that analysis is carried out through precise and relevant logical steps and is of great support to further the understanding of the evaluation and its findings. The cause and effect links between intervention and consequences are clearly explained and limitations are highlighted (Vol. 2b p. 217-218).

Comparisons are made, in particular for before/after and with/without (on data related to access or quality indicators, and for example the funding of a Handicap International programme in Niger). The report systematically refers to the annexes which further develop the basis for the analysis and shows that the statements made in the report are well supported.

6. Credible findings: GOOD

Findings are the logical result of data analysis and interpretation. They are balanced and clearly presented. The combination of evidence provided by the different data collection tools is well managed and balanced. The perceptions of actors and beneficiaries, collected through the surveys, interviews and focus groups, are presented in almost every EQ, adding to the credibility of the findings. Both internal validity and external validity are satisfactory and supported by the great number of information collected through independent sources and in different contexts.
7. Validity of the conclusions:  GOOD

Conclusions are clear, logical consequences of findings and easily linkable with the facts and data recorded in the evaluation. Concepts are well summarized and the organization of conclusions in three clusters makes them even more straightforward. Limitations to conclusions are not assessed.

8. Usefulness of the recommendations:  GOOD

Recommendations are clearly linked to conclusions and focus on the main key findings of the evaluation. They are substantiated by findings and presented in a very clear and structured way (organized in three separate clusters). Recommendations are operational in that they provide practical options and steps to be taken to address the challenges highlighted through the evaluation. However, a more 'strategic' and 'synthetic' approach could have made recommendations more directly applicable. Responsibilities to act and recommendations are clearly prioritized in terms of importance and urgency (p. 89 Main report).

9. Clearly reported:  VERY GOOD

The report is clear and well structured; it offers complete and exhaustive information on the scope of the evaluation, the design and its results. The report is easy to read and rich of details that will be of great use for a more technical audience. The summary is clear and presents the main conclusions and recommendations in a balanced and unbiased way. Annexes are extremely useful and provide a very thorough back up to the report, which remains, nevertheless, self standing.