Checklist for Assessing USAID Evaluation Reports

Checklist for Reviewing Evaluation Reports

High quality, evidence-based evaluation reports with a clear focus on decision-making for USAID and other key partners are critical for improving USAID's development effectiveness. USAID and evaluators need to work collaboratively to ensure high quality evaluations using clear and specific standards. Use the Evaluation Report Checklist to review and strengthen draft evaluation reports. The fifteen (15) items in bold are critical factors that should be addressed in early drafts of the evaluation report. As the report is finalized we recommend that you assess it against all seventy six (76) factors to ensure high technical quality, a strong executive summary, and the targeting of recommendations for decision-making purposes. One of USAID's main responsibilities with evaluations is to ensure that they are broadly disseminated--and actively communicated--for learning, program improvement and accountability purposes. Make sure to pay attention not only to the technical quality of evaluations but to promoting their use and impact.

EVALUATION REPORT CHECKLIST - V1.0

Title of Study Being Reviewed: ______________________________

Main Implementer(s): ______________________________

Reviewer: ______________________________

Date of Review: ______________________________

GOOD PRACTICE ELEMENTS OF AN EVALUATION REPORT[i]
Keyed to USAID's 2011 Evaluation Policy

EVALUATION REVIEW FACTOR	1	2	3	4	5	Reviewer Comments
STRUCTURE OF THE REPORT						
1. Does the evaluation report have a cover sheet attached indicating the type of evaluation conducted (e.g. performance evaluation or impact evaluation) and general design?						
2. If a performance evaluation, does the evaluation report focus on descriptive and normative evaluation questions?						
3. If the evaluation report uses the term "impact evaluation," is it defined as measuring the change in a development outcome that is attributable to a defined intervention (i.e. impact evaluations are based on models of cause and effect and require a credible and rigorously defined counterfactual)?						
4. Regardless of the type of evaluation, does the evaluation report reflect use of sound social science methods?						
5. Does the report have a Table of Contents (TOC)?						
6. Do Lists of Figures and Tables follow the TOC?						
7. Does the report have a Glossary of Terms?						
7.1 Are abbreviations limited to the essential?						
8. Is the date of the report given?						
9. Does the body of the report adhere to the 20 page guide?						
10. Is the report well-organized (each topic is clearly delineated, subheadings used for easy reading)?						
11. Does the report's presentation highlight important information in ways that capture the reader's attention?						
12. Is the report well written (clear sentences, reasonable length paragraphs, no typos, acceptable for dissemination to potential users)?						
13. Does the evaluation report focus on the essential issues concerning the key questions, and eliminate the "nice to know", but not essential information?						

Key: 1 = Not addressed; 2 = Poorly/partially addressed; 3 = Adequately addressed; 4 = More than adequately addressed; 5 = Exemplar.

EVALUATION REVIEW FACTOR	1	2	3	4	5	Reviewer Comments
14. Does the evaluation report discuss any issues of conflict of interest, including the lack thereof?						
15. As applicable, does the evaluation report include statements regarding any significant unresolved differences of opinion on the part of funders, implementers and/or members of the evaluation team?						
EXECUTIVE SUMMARY						
16. Does the evaluation report begin with a 3- to 5-page stand-alone summary of the purpose, background of the project, main evaluation questions, methods, findings, conclusions, recommendations and lessons learned (if applicable) of the evaluation?						
17. Does the Executive Summary concisely state the main points of the evaluation?						
18. Does the Executive Summary follow the rule of only saying what the evaluation itself says and not introducing new material?						
INTRODUCTION						
19. Does the report introduction adequately describe the project?						
19.1. Does the introduction explain the problem/opportunity the project was trying to address?						
19.2. Does the introduction show where the project was implemented (physical location) through a map?						
19.3. Does the introduction explain when the project was implemented?						
19.4. Are the “theory of change” or development hypotheses that underlie the project explained? (Does the report specify the project’s inputs, direct results (outputs), and higher level outcomes and impacts, so that the reader understands the logical structure of the project and what it was supposed to accomplish?)						
19.5. Does the report identify assumptions underlying the project?						
19.6. Does the report include sufficient local and global contextual information so that the external validity and relevance of the evaluation can be assessed?						
19.7. Does the evaluation report identify and describe any critical competitors to the project that functioned at the same time and in the project’s environment?						
19.8. Is USAID’s level of investment in the project stated?						
19.9. Does the evaluation report describe the project components funded by implementing partners and the amount of funding?						
20. Is the purpose of the evaluation clearly stated?						
21. Is the amount of USAID funding for the evaluation indicated?						
22. Are all other sources of funding for the evaluation indicated as well as the amounts?						
23. Does the report identify the evaluation team members and any partners in the evaluation?						
24. Is there a clear statement of how the evaluation will be used and who the intended users are?						
25. Are the priority evaluation questions presented in the introduction?						
26. Does the evaluation address all evaluation questions included in the Statement of Work (SOW)?						
26.1. Are any modifications to the SOW, whether in technical requirements, evaluation questions, evaluation team composition, methodology or timeline indicated in the report?						
26.2. Is the SOW presented as an annex?						
26.3. If so, does the annex include the rationale for any change with the written sign-offs on the changes by the technical officer?						

Key: 1 = Not addressed; 2 = Poorly/partially addressed; 3 = Adequately addressed; 4 = More than adequately addressed; 5 = Exemplar.

EVALUATION REVIEW FACTOR	1	2	3	4	5	Reviewer Comments
SCOPE AND METHODOLOGY						
27. Does the report provide a clear description of the evaluation's design?						
27.1. Is a design matrix or similar written tool presented in an annex that shows *for each question/subquestion* the measure(s) or indicator(s) used to address it, the source(s) of the information, the type of evaluation design, type of sampling if used, data collection instrument(s) used, and the data analysis plan?						
28. Does the report state the period over which the evaluation was conducted?						
29. Does the report state the project time span (reference period) covered by the evaluation?						
30. Does the evaluation report indicate the nature and extent of consultation on the evaluation design with in-country partners and beneficiaries?						
31. Does the evaluation report indicate the nature and extent of participation by national counterparts and evaluators in the design and conduct of the evaluation?						
32. Does the report address each key question around which the evaluation was designed?						
33. Is at least one of the evaluation questions directly related to gender analysis of outcomes and impacts?						
33.1. Are data sex-disaggregated?						
34. In answering the questions, does the report appropriately use comparisons made against baseline data?						
35. If the evaluation is expected to influence resource allocation, does it include information on the cost structure and scalability of the intervention, as well as its effectiveness?						
35.1. As appropriate, does the report include financial data that permits computation of unit costs and analysis of cost structure?						
36. Is there a clear description of the evaluation's data collection methods (summarized in the text with the full description presented in an annex)?						
36.1. Are all tools (questionnaires, checklists, discussion guides, and other data collection instruments) used in the evaluation provided in an annex?						
36.2. Does the evaluation report include information, as appropriate, on the pilot testing of data collection instruments?						
36.3. Does the evaluation report include information, as appropriate, on the training of data collectors?						
37. Are all sources of information properly identified and listed in an annex?						
38. Does the evaluation report contain a section describing the limitations associated with the evaluation methodology (e.g. selection bias, recall bias, unobservable differences between comparator groups, small samples, only went to villages near the road, implementer insisted on picking who the team met with, etc)?						
39. Does the evaluation report indicate the evaluation methodology took into account the time, budget, and other practical considerations for the evaluation such as minimizing disruption and data burden?						
40. Does the report have sufficient information to determine if the evaluation team had the appropriate methodological and subject matter expertise to conduct the evaluation as designed?						
41. If an impact evaluation was designed and conducted, does the evaluation report indicate that experimental methods were used to generate the strongest evidence? Or does the report indicate that alternative methods for assessing impact were utilized and present the reasons why random assignment strategies were not feasible?						
42. Does the evaluation report reflect the application and use to the maximum						

Key: 1 = Not addressed; 2 = Poorly/partially addressed; 3 = Adequately addressed; 4 = More than adequately addressed; 5 = Exemplar.

EVALUATION REVIEW FACTOR	1	2	3	4	5	Reviewer Comments
extent possible of social science methods and tools that reduce the need for evaluator-specific judgments?						
43. Does the evaluation scope and methodology section address generalizability of the findings?						
ANALYSIS						
44. Are percentages, ratios, cross-tabulations, rather than raw data presented, as appropriate?						
45. When percentages are given, does the report always indicate the number of cases used to calculate the percentage?						
45.1. Is use of percentages avoided when the number of cases is small (<10)?						
46. Are whole numbers used or rounding-off numbers to 1 or 2 digits?						
47. Are pictures used to good effect?						
47.1. Relevant to the content						
47.2. Called out in the text and placed near the call-out						
48. Are charts and graphs used to present or summarize data, where relevant?						
48.1. Are the graphics easy to read and simple enough to communicate the message without much text?						
48.2. Are they consistently numbered and titled?						
48.3. Are they clearly labeled (axis, legend, etc.)						
48.4. Is the source of the data identified?						
48.5. Are they called out in the text and correctly placed near the call-out?						
48.6. Are the scales honest (proportional and not misleading by virtue of being “blown-up”)?						
FINDINGS						
49. Are FINDINGS specific, concise and supported by strong quantitative and qualitative evidence?						
49.1. As appropriate, does the report indicate confirmatory evidence for FINDINGS from multiple sources, data collection methods, and analytic procedures?						
50. Are adequate data provided to address the validity of the “theory of change” or development hypothesis underlying the project, i.e., cause and effect relationships?						
51. Are alternative explanations of any observed results discussed, if found?						
52. Are unplanned results the team discovered adequately described?						
53. Are opinions, conclusions, and recommendations kept out of the description of FINDINGS?						
CONCLUSIONS						
54. Is there a clear distinction between CONCLUSIONS and FINDINGS?						
55. Is every CONCLUSION in the report supported by a specific or clearly defined set of FINDINGS?						
56. Are the CONCLUSIONS credible, given the FINDINGS the report presents?						
57. Can the reader tell what CONCLUSIONS the evaluation team reached on each evaluation question?						
RECOMMENDATIONS						
58. Are RECOMMENDATIONS separated from CONCLUSIONS? (Are they highlighted, presented in a separate section or otherwise marked so that the reader sees them as being distinct?)						
59. Are all RECOMMENDATIONS supported by a specific or clearly defined set of FINDINGS and CONCLUSIONS? (Clearly derived from what the evaluation team learned?)						
60. Are the RECOMMENDATIONS practical and specific?						

Key: 1 = Not addressed; 2 = Poorly/partially addressed; 3 = Adequately addressed; 4 = More than adequately addressed; 5 = Exemplar.

EVALUATION REVIEW FACTOR	1	2	3	4	5	Reviewer Comments
61. Are the RECOMMENDATIONS responsive to the purpose of the evaluation?						
62. Are the RECOMMENDATIONS action-oriented?						
63. Is it clear who is responsible for each action?						
64. Are the RECOMMENDATIONS limited/grouped into a reasonable number?						
LESSONS LEARNED						
65. Did this evaluation include lessons that would be useful for future projects or programs, on the same thematic or in the same country, etc.?						
66. Are the LESSONS LEARNED highlighted and presented in a clear way?						
67. Does the report indicate who the lessons are for? (e.g., project implementation team, future project, USAID and implementing partners, etc.)						
BOTTOM LINE						
68. Does the evaluation report give the appearance of a thoughtful, evidence-based, and well organized effort to objectively evaluate what worked in the project, what did not and why?						
69. As applicable, does the evaluation report include statements regarding any significant unresolved differences of opinion on the part of funders, implementers and/or members of the evaluation team?						
70. Is the evaluation report structured in a way that will promote its utilization?						
71. Does the evaluation report explicitly link the evaluation questions to specific future decisions to be made by USAID leadership, partner governments and/or other key stakeholders?						
72. Does the evaluation report convey the sense that the evaluation was undertaken in a manner to ensure credibility, objectivity, transparency, and the generation of high quality information and knowledge?						

REPORT DISSEMINATION						
73. Have all evaluation team members signed a statement attesting to a lack of conflict of interest, or describing and existing conflict of interest relative to the project being evaluated?						
74. Was the Report Submitted to the Development Experience Clearing House (DEC)?						
75. Has a dissemination plan been developed for this report?						
76. Is the report widely shared to interested stakeholders?						

DEFINITIONS:

Performance evaluation: focuses on descriptive and normative questions: what a particular project or program has achieved (either at an intermediate point in execution or at the conclusion of an implementation period); how it is being implemented; how it is perceived and valued; whether expected results are occurring; and other questions that are pertinent to program design, management and operational decision making. Performance evaluations often incorporate before-after comparisons, but generally lack a rigorously defined counterfactual.

Impact evaluation: measures the change in a development outcome that is attributable to a defined intervention; impact evaluations are based on models of cause and effect and require a credible and rigorously defined counterfactual to control for factors other than the intervention that might account for the observed change. Impact evaluations in which comparisons are made between beneficiaries that are randomly assigned to either a ―treatment‖ or a ―control group provide the strongest evidence of a relationship between the intervention under study and the outcome measured.

Theory of change: A tool to design and evaluate social change initiatives. It is a blueprint of the building blocks needed to achieve long-term goals of a social change initiative.

Key: 1 = Not addressed; 2 = Poorly/partially addressed; 3 = Adequately addressed; 4 = More than adequately addressed; 5 = Exemplar.

Development Hypothesis: Identifies causal linkages between USAID actions and the intended Strategic Objective (highest level result).

External Validity: The degree to which findings, conclusions, and recommendations produced by an evaluation are applicable to other settings and contexts.

Findings: Empirical facts collected during the evaluation

Conclusions: Interpretations and judgments based on the findings

Recommendations: Proposed actions for management.

---

[i] In addition to the USAID 2011 Evaluation Policy, good practices in evaluation reporting have also been drawn from:
Morra Imas, Linda and Ray C. Rist. 2009. *The Road to Results: Designing and Conducting Effective Development Evaluations.* Washington, DC.: The World Bank.
Scriven, Michael. 2005. *Key Evaluation Checklist.*
Stufflebeam, Daniel L. 1999. *Program Evaluations Metaevaluation Checklist*.