



## CSO Evaluability Assessment Checklist: Working Draft

---

### What Is an Evaluability Assessment?

An Evaluability Assessment (EA) is a review, generally conducted prior to an evaluation, to determine an evaluation's return on investment. By reviewing the plausibility and feasibility of a project's design, the quality of available information, and the project's context, an EA can determine the feasibility and learning utility of a prospective evaluation or evaluation question. Based on this assessment of the costs and benefits of conducting an evaluation, CSO can then decide if that would be a good investment of resources.

### Why Conduct an EA?

An EA can prevent us from pursuing evaluations with low returns on investment; in other words an evaluation that costs CSO more than what it gains in terms of learning or accountability. We can use EA results to write evaluation scopes of work that set up useful evaluations. In some cases, an EA is able to identify valuable learning for the bureau without the additional costs of data collection associated with an evaluation. Finally, when conducted early in a program's life-cycle, an EA can help identify and correct program weaknesses, increasing the likelihood the program will achieve that desired results.

### Who Conducts an EA?

EAs can be conducted by individuals internal, external or a combination of both who have strong program design, monitoring and evaluation competencies. These individuals can go on to form part or all of the evaluation team, if the bureau decides to pursue an evaluation, or can conclude their involvement after the EA.

### When Is the Right Time for an EA?

There are two stages in the program life cycle that an EA is commonly implemented; just prior to deciding whether to evaluate a program or immediately following program design to identify any weaknesses in program analysis, design, and/or monitoring planning. Some of the EA checklist questions would need to be adapted slightly if used early in a program's life cycle. For example, immediately following a project's design, a project would not yet have collected any data, so the EA team would review the monitoring plan to assess if it offers the right information from the right sources, using appropriate methods, at the right time. At that stage, the EA team would assess whether indicators are specific, measurable, accurate, reliable, and time bound (SMART). Conversely, an EA to inform the decision about whether to evaluate would not need to assess the indicators but would need to determine whether the results data were valid.

### What Is an EA Checklist?

Conducting an EA involves gathering information about a project and systematically analyzing it to reach recommendations about whether to evaluate and/or modify the

project. An EA checklist systematizes that analysis and offers the basis against which to assess the return on investment of an evaluation.

### What Else Goes into an EA Beyond Using a Checklist?

There are multiple steps in conducting an EA. The U.S. Agency for International Development's guidance, for example, lists five steps, starting with clarifying the EA's purpose and ending with making recommendations about a program's readiness for evaluation. The UK's Department for International Development commissioned a synthesis of EA literature in which Rick Davies summarizes additional approaches to the overall process.

The CSO EA Checklist does not provide guidance on the full EA process. In the course of conducting our pilot EA, however, we identified a gap in the existing guidance: how to move from answers to checklist questions to recommendations about whether, what, and how to evaluate. To fill this, we used a return-on-investment (ROI) approach to analyze the data applying the EA checklist generates. A ROI approach seeks to understand a potential evaluation's costs and benefits, i.e., the money, time, political capital and other resources that would be required to complete an evaluation and the utility of the learning it could generate.

Given the lack of other guidance on this approach, it is worth briefly noting a few key points:

- There are certain minimum threshold criteria that, if not met, mean no evaluation should proceed. These criteria are: safety; timing (i.e., there is ongoing/future work an evaluation could inform); and availability of core program information.
- Most criteria, however, do not automatically suggest an evaluation either should or should not proceed. The implication of not meeting a criterion is context specific. For example, the lack of a program design need not turn off an evaluation if the program team is willing to retroactively create a theory of change and/or is open to goal free evaluation approaches. The EA team's job, therefore, is not simply to document whether or not there is a program design but rather what the lack of that design means it would cost to complete an evaluation and how useful the potential learning would be.
- A ROI analysis leads to more nuanced recommendations than just whether or not to evaluate a program. It also looks at which evaluation questions are likely to have the highest return on investment and which evaluation approaches would be most appropriate.

For a list of documents generally needed to conduct an EA, such as a baseline report, see Annex C.

### How Did We Develop This Checklist?

CSO started with the EA checklist compiled by Rick Davies for the UK Department for International Development in 2013. We expanded the content to be more accessible to different audiences. Many existing EA resources are written for evaluators. This checklist

is intended to be a tool for program staff as well so it provides greater explanation of each criterion. Then we filtered the checklist (i.e. criteria, framing questions, and characteristics) through the lens of CSO's conflict mitigation mandate, place within the U.S. government, and preference towards Utilization Focused Evaluation. This checklist, therefore, is CSO-specific, but easily adaptable by others in and beyond the U.S. government. Further, this checklist assumes the actors involved will be willing to engage in a full range of evaluation approaches.

Note: CSO used an earlier version of this checklist to conduct an EA and made revisions based on that experience. CSO considers the checklist to be a working draft until the bureau uses it to conduct at least one additional EA.

For guidance on terminology see Appendix A: Terms.

#### How Does One Read the Checklist?

The checklist is broken out into four sections. Each section has the same structure. The left column lists the criteria that make up the checklist, phrased as a question to help provide clarity to the intent. The right column lists characteristics a project fully meeting each criterion would exhibit.

The first section, minimum threshold, lists three criteria programs must be met to make an evaluation a worthwhile use of the time. In other words, there is no possible evaluation that could offer a positive return on investment if a program does not meet these three criteria. The minimum threshold criteria are meant to be applied at a general level. For instance, can participants be in contact with individuals identified with this project or a foreign agency? This question does not need to be answered for every possible grouping or faction in the context but for the average participant. The criteria will be revisited later in the process in the context of specific evaluation questions and will receive a more detailed application at this time.

The next three sections are: (A) Project Design; (B) Information Availability; and (C) Institutional Context, with subsections on practicality and utility.

| <b>0. Minimum Thresholds</b> |   |
|------------------------------|---|
| <b>Criteria</b>              | <b>Characteristics</b>  |
| Safety                       | <ul style="list-style-type: none"> <li>• CSO and, if applicable, embassy personnel deem the physical security risks for the evaluation team and participants in the evaluation as acceptable.</li> <li>• CSO concludes conducting an evaluation will not negatively impact the project or key relationships.</li> </ul> |
| Timing                       | <ul style="list-style-type: none"> <li>• There is ongoing or future work and/or decisions that the evaluation can inform.</li> </ul>  |
| Core Information             | <ul style="list-style-type: none"> <li>• Sufficient critical project documentation is available including documents related to context analysis, program design, program implementation, monitoring, evaluation, and/or learning. The ability to contact program beneficiaries is particularly important.</li> </ul>    |

| <b>A. Project Design</b>  |  |
|---|--|
| <b>Criteria</b>   | <b>Characteristics</b>   |
| <p><b>A1. Clarity</b><br/>How clearly stated are the changes the project sought to achieve and the activities it undertook?</p> | <p><i>Projects should adapt to the context to optimize effectiveness and relevance throughout the project. Clarity is valuable at each stage in the evolution.</i></p> <ul style="list-style-type: none"> <li>• The most significant change (i.e. goal) this project intended to achieve, and each of its evolutions over time, is identifiable.</li> <li>• The original objectives and all of their evolutions over time are clear change statements (i.e. not activities).</li> <li>• Project personnel and documents use similar language to describe the goal and objectives (and their evolutions).</li> <li>• The scale of the work effort is clear, including: <ul style="list-style-type: none"> <li>○ The quantity of work (e.g., 15 trainings for 35-30 participants each);</li> <li>○ The number of work sites or the geographic coverage (e.g. X village); and</li> <li>○ The target of the work (e.g. rural school teachers).</li> <li>○ The number of staff and partners involved.</li> </ul> </li> <li>• Adaptations to the project are documented and the rationale explained with supporting evidence, e.g. monitoring data.</li> </ul> |
| <p><b>A2. Plausible</b><br/>How likely is it the Theory of Change (TOC) will create change on the ground?</p>                   | <ul style="list-style-type: none"> <li>• There is an overarching strategy (e.g., TOC) that realistically connects the program to peace writ large.</li> <li>• There is a clearly articulated explanation behind how CSO's work will catalyze the goal.</li> <li>• The project goal is <i>feasible</i>. It could be achieved within the project lifespan, activities, output dosage, and resources (budget, staffing) in this context.</li> <li>• The TOC is justified through evidence or assumptions. Evidence of effectiveness, best practices, research findings, or lessons learned are acceptable, as is an explicit articulation of the assumptions explaining how the change process would unfold.</li> <li>• The TOC takes into account the needs of different subgroups within the target population, e.g.</li> </ul>   |

|  |  |
|--|--|
|  | <p>women, youth, or the disabled.</p> <ul style="list-style-type: none"> <li>• Project documents (e.g., the design/proposal, M&amp;E plans, work plans, progress reports, etc.) and key partners (e.g. embassy staff, CSO, and implementing partners) are consistent in the way they describe the (TOC) at each evolution.</li> <li>• There is an explicit analysis of who a project’s influencers and spoilers would be and how they would affect the project.</li> </ul> |
| <p><b>A3. Relevant</b><br/>Is the project goal directly pertinent to the conflict’s key driving factors?</p> | <ul style="list-style-type: none"> <li>• The conflict analysis identifies the conflict’s <i>key</i> driving factors (i.e., it is not a general description of the history of the conflict).</li> <li>• The importance of the project goal is explained and clearly derived from the conflict analysis.</li> <li>• The conflict analysis has been updated and the project goal remains relevant or has been adapted.</li> </ul>   |
| <p><b>A4. Do No Harm</b><br/>Is the project conflict sensitive?</p>  | <ul style="list-style-type: none"> <li>• Project documents include a conflict-sensitivity analysis that identifies how the project could exacerbate or positively contribute to the conflict.</li> <li>• The project design components that reflect/address conflict sensitivity clearly reflect this analysis.</li> </ul>   |
| <p><b>A5. Programs</b><br/>Does the design factor in other programs?</p>                                     | <ul style="list-style-type: none"> <li>• Other projects with similar goals and target populations/participants are documented and the relationship with this project is clear.</li> </ul>  |
| <p><b>A6. Alignment</b><br/>Does the project align to U.S. government policies?</p>                          | <ul style="list-style-type: none"> <li>• The U.S. government policy this project supports is clearly articulated.</li> <li>• If the policy changed, the program’s evolution to remain policy-relevant is clearly documented.</li> <li>• Activities mandated or limited by U.S. policy are clearly noted.</li> </ul>  |

| <b>B. Information Availability</b>                                     |   |
|--|---|
| <b>Criteria</b>  | <b>Characteristics</b>  |
| <p><b>B1. Core Information</b><br/>Is the core material available?</p> | <ul style="list-style-type: none"> <li>• Names, roles, and contact information of implementing actors are available, including CSO staff, implementing partners, consultants (e.g. evaluators), and U.S. donor partners within and outside of the U.S. government.</li> <li>• The intended and actual participants in project activities are documented. There are records of who was involved in what project activities when and their contact information is available.</li> </ul> |

|  |   |
|--|---|
|  | <ul style="list-style-type: none"> <li>• The classification of information makes it available to an evaluation team.</li> </ul>   |
| <b>B2. Baseline</b><br>Do baseline measures exist?   | <ul style="list-style-type: none"> <li>• The baseline determined a starting point of key changes before the program began or as quickly thereafter as possible that was directly derived from the project design and assessed with SMART indicators, where appropriate.</li> <li>• The baseline raw data (e.g. individual survey responses) is available in an electronic form.</li> <li>• Baseline data has key disaggregation. Key will need to be determined by project and context.</li> <li>• The methodology e.g., sampling, data collection instruments was “good enough” to generate credible data.</li> <li>• If baseline data is in the form of national or subnational statistics, time-series data is available for pre-project years.</li> </ul> |
| <b>B3. Control Group</b><br>Is there data from a control/comparison group?                   | <ul style="list-style-type: none"> <li>• It is clear how the characteristics of the control/comparison group compare to the intervention group.</li> <li>• The members of the control/comparison group are identifiable and contactable.</li> <li>• The raw data is available in electronic form.</li> <li>• Data has key disaggregation. Key will need to be determined by project and context.</li> <li>• The number of times CSO (or designees) collected on the status of the control/comparison group is documented.</li> <li>• The frequency of data collection from the control/comparison group is sufficient to determine any meaningful differences from the intervention group.</li> </ul>   |
| <b>B4. Results</b><br>Is results data available (including intended and unintended results)? | <ul style="list-style-type: none"> <li>• Data (generally via SMART indicators) documenting progress (or lack of) towards results is available electronically.</li> <li>• There is no significant missing data.</li> <li>• Data collection methods were “good enough”</li> <li>• Appropriate data disaggregation exists.</li> <li>• The program scanned for and documented unintended results, including conflict-sensitivity issues.</li> </ul>   |
| <b>B5. Context</b><br>Is context monitoring information available?                           | <ul style="list-style-type: none"> <li>• Data documenting whether or not the contextual assumptions remain valid is available for the life of the program.</li> <li>• Monitoring data on influencers’ and spoilers’ behaviors is available.</li> </ul>  |
| <b>B6. Performance Standards</b><br>Are there performance                                    | <ul style="list-style-type: none"> <li>• CSO staff and implementing partners agree on what constitutes excellence in implementation processes for the program, e.g. planning, facilitation, peace journalism, dialogue that is tailored to the context.</li> </ul>  |

|   |   |
|---|---|
| standards for relevant tasks?   | <ul style="list-style-type: none"> <li>• Raw data, that has been appropriately disaggregated, is accessible against the performance standards.</li> </ul>   |
| <p><b>B7. Evaluations</b></p> <p>Is the information from past evaluation processes available?</p> | <ul style="list-style-type: none"> <li>• Evaluative products (e.g. reports) are available that include the methodology and results of the evaluative processes (e.g. after action review, evaluations).</li> <li>• The raw data is available.</li> <li>• The methodology (e.g., from sampling to data collection instruments) was appropriate and reliable.</li> <li>• Data has key disaggregation.</li> <li>• The evaluand felt it was a useful process and valuable product.</li> </ul> |



| <b>C. Institutional Context: Practicality</b>  |   |
|--|---|
| <b>Criteria</b>  | <b>Characteristics</b>  |
| <p><b>C1. Accessibility</b><br/>Will the evaluation team be able to access key people?</p>                                     | <ul style="list-style-type: none"> <li>• Relevant people are likely to agree to participate in data collection (e.g. they are not too fearful, suspicious, busy, or prevented from participating by someone in power).</li> <li>• No significant displacement of individuals necessary to an evaluation has occurred since CSO collected participant contact information.</li> <li>• Technology is an option to reach participants. This means there's electricity, mobile connectivity, and the government has not shut down the internet.</li> <li>• The embassy is comfortable with an evaluation team accessing those key individuals where there may be political or other sensitivities, such as government officials, militias, opposition groups, political organizations, and/or armed groups.</li> <li>• Program staff and partners will be present at the time of the evaluation.</li> <li>• Weather will not limit the evaluation process, e.g., the evaluation is not being planned during a rainy season that would prevent the evaluation team from visiting key sites.</li> <li>• No significant events that would disrupt or distort (positively or negatively) an evaluation process (e.g., data collection or ability to bring people together) will occur during the evaluation, such as an election or anniversary of a massacre.</li> </ul> |
| <p><b>C2. Resources</b><br/>Are there sufficient resources to do the evaluation?</p>   | <ul style="list-style-type: none"> <li>• There is sufficient total time and time in-country available to conduct an evaluation.</li> <li>• There is sufficient funding available for the relevant team and methodological options (e.g. household surveys or social media analysis).</li> <li>• Note: determinations of sufficient time and funding must take into account how common the skills necessary for the evaluation are.</li> </ul>   |
| <p><b>C3. Coordination</b><br/>Have those who need to be involved to make the evaluation a success been involved/notified?</p> | <ul style="list-style-type: none"> <li>• CSO has identified any donors, U.S. government agencies, and NGOs, relevant to the evaluation and the level of engagement necessary, ranging from joint evaluation to courtesy notification.</li> <li>• The embassy, regional bureau, relevant functional bureaus, and CSO leadership (at least at the office level) support an evaluation process.</li> </ul>   |
| <b>C. Institutional Context: Utility</b>   |   |
| <b>Criteria</b>  | <b>Characteristics</b>  |

|   |   |
|---|---|
| <p><b>C4. Timing</b><br/>Is the timing right for this evaluation?</p>   | <ul style="list-style-type: none"> <li>• There is sufficient time for the evaluation to meet the deadlines necessary to be useful.</li> <li>• The timing of the evaluation occurs at a time in the project implementation that will allow the evaluation question to be answered (e.g., there has been sufficient work that change is possible. )</li> </ul>  |
| <p><b>C5. Catalysts</b><br/>Will the primary users be engaged in the evaluation?</p>                          | <ul style="list-style-type: none"> <li>• CSO has identified primary users of the evaluation who are acceptable to all possible users.</li> <li>• The primary users can be involved in scoping the evaluation and are willing to participate in an evaluation process.</li> </ul>  |
| <p><b>C6. Learning</b><br/>Have the primary users prioritized their desired learning?</p>                     | <ul style="list-style-type: none"> <li>• The primary users have prioritized the evaluation questions based on the learning potential for specific users.</li> </ul>   |
| <p><b>C7. Process Expectations</b><br/>Have the primary users conveyed their expectations to the EA team?</p> | <ul style="list-style-type: none"> <li>• Primary users have expressed their expectations of the evaluation process. (Note: where supervisors' views influence user engagement in the evaluation process, those should be included in the expectations users express.)</li> <li>• Primary users view a range of evaluation approaches and data collection methodologies as credible.</li> <li>• A realistic expectation exists among users about what they will do with evaluation results.</li> </ul> |
| <p><b>C8. Ethics</b><br/>Are the ethical issues manageable?</p>   | <ul style="list-style-type: none"> <li>• The EA team, in consultation with others, is able to identify risks and benefits to evaluation participants and implementing partners, as well as mitigation strategies to counter possible risks.</li> </ul>  |
| <p><b>C9. Challenges</b><br/>Are the challenges that will face an evaluation team manageable?</p>             | <p>As challenges vary, here are some common questions to consider:</p> <ul style="list-style-type: none"> <li>• Will primary users be able to accept, apply, and/or constructively respond to negative findings?</li> <li>• Have previous evaluation experiences prejudiced the primary users likely participation?</li> <li>• Will classification of information make finding skilled evaluators difficult?</li> <li>• Will participants speak candidly to the evaluation team?</li> </ul>           |

## Appendix A: Terms

- Assumptions: the rationale or underlying beliefs that explain the relationship between objectives.
- Baseline: the determination of the initial measure or status of the elements a project seeks to change.
- Best practice: approaches or tactics that are perceived to be effective based on accumulated experience.
- Change: a difference in the content, form, or functioning of something in a context that is separate from an implementing actor's actions.
- Conflict sensitivity: the process of minimizing negative impacts and maximizing positive impacts of an intervention on conflict.
- Data disaggregation: data that has been broken out into component parts, e.g. by age, gender, religion.
- Evaluability: the degree of difficulty associated with evaluating a project.
- Evaluand: subject of an evaluation.
- Evaluation manager: the CSO staff person responsible for communicating with the EA/evaluation team and the evaluand, preserving the EA/evaluation's independence, and ensuring that the EA/evaluation team receives the information from CSO necessary to produce a high-quality product.
- Evaluative process: any intentional process that uses the appropriate amount of data to understand the value and quality (i.e., merit and worth) of a program or policy. Data may be a significant or a minor element of the process.
- Evidence: data or information that proves or disproves an assertion.
- Goal: the most significant change a project sets out to achieve in a set period of time.
- Influencer: someone within the context who mobilizes for or mitigates changes the project is seeking.
- Output: the immediate, generally tangible, result of the activities; never a change nor the activity itself.
- Output dose: the amount of outputs that are generated by the activities.
- Technology: any use of information communication capability in the evaluation process, e.g. mobile data collection, crowdsourcing, social media monitoring.
- Primary user: the principle audience of the evaluation, who must be in a position to take action on the evaluation's recommendations.
- Relevant: a project goal that responds to a significant conflict prevention/mitigation need, driver, or cause.
- Reliable: data collection will produce the same results over time if conducted by different people.
- SMART: an acronym for: specific, measurable, accurate, reliable, and time bound, which are characteristics of an effective indicator.
- Spoiler: any actor who will work to undermine the success of a project.
- Time-series data: a sequence of data points collected over a time interval.
- Valid: the data collection method measures what it is intended to measure.

## **Appendix B: Sources**

Davies, R., 2013. *Planning Evaluability Assessments: A Synthesis of the Literature with Recommendations*. Report of a Study Commissioned by the Department for International Development. Available at:

[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/248656/wp40-planning-eval-assessments.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/248656/wp40-planning-eval-assessments.pdf)

ILO Evaluation Unit, Using the Evaluability Assessment Toolkit, Guidance Note 11, 2011

Reimann, C., Evaluability Assessments in Peacebuilding Projects, CDA Working Paper, 2012

[Scharbatke-Church, C., and Rogers, M.M. Evaluability Assessment draft chapter, \*Designing for Results Second Edition\*, forthcoming 2016.](#)

USAID, Evaluation, Conducting an Evaluability Assessment for USAID Evaluations, An Evaluation Resource, 2015. Available at:

[https://usaidlearninglab.org/sites/default/files/resource/files/conducting\\_an\\_evaluability\\_assessment\\_july2015.pdf](https://usaidlearninglab.org/sites/default/files/resource/files/conducting_an_evaluability_assessment_july2015.pdf)

## Appendix C: Common Documents

Every EA is different, however there are some common documents that typically are useful when conducting an EA. These include:

1. Context/conflict analysis
2. Actor/Stakeholder analysis (maybe part of conflict analysis)
3. Program descriptions:
  - a. Classic program proposal documents
  - b. Hand over/transition reports
  - c. Workplans
  - d. Products of the program's activities, if appropriate e.g. radio spot, report, etc.
4. Baseline report
5. Baseline instrument
6. Raw data files from baseline
7. Monitoring plan
8. Monitoring reports e.g. quarterly reports
9. Monitoring data collection instruments
10. Raw data files from monitoring
11. Internal reviews or reflection session notes e.g. After Action Review
12. Contact lists, participant list, partners lists; all with contact information
13. Security situation updates
14. Evaluation Scope of Work
15. Evaluation Reports
16. Evaluative process raw data
17. Organigram with role descriptions

This EA checklist is part of CSO's commitment to learning and pursuing useful evaluations. As part of this effort, CSO is working with Besa's Cheyanne Scharbatke-Church.