Chapter 9

STAGE TWO

EVALUATION MANAGEMENT

This chapter contains:

1. Developing the Terms of Reference
2. The Evaluation Plan
3. Frequently Asked Questions about Working with External Evaluators
4. Strategies for Overcoming Common Evaluation Pitfalls

INTRODUCTION

Evaluation Management involves the implementation of decisions made in the preparation stage. The process begins with the creation of the terms of reference (TOR), which generally takes place four to six months prior to the evaluation date. Developed by the project team, the TOR is essentially a guide to the evaluation and, as such, needs to be well thought through.

After completing the TOR, the next step in the evaluation management stage is to develop the evaluation plan. This can be done by the project team or by the evaluators and always should have input from both parties. As the second step in the evaluation management process, the evaluation plan operationalizes the decisions made in the TOR. This requires additional decisions related to the means of verification, data sources and targets, location of data collection, conflict considerations, means of analysis, and timing.

Evaluation Management involves the implementation of decisions made in the preparation stage.

Working with external evaluators also requires some thought to ensure a successful experience. This chapter walks through a number of the most frequently asked questions about external evaluators starting with the recruitment process through to what to do with the evaluator's primary data after the evaluation is completed.

The Managing Evaluations chapter concludes by offering some strategies for overcoming common evaluation pitfalls.

I. DEVELOPING THE TERMS OF REFERENCE

What are the terms of reference?

The first step in evaluation preparation is the development of the terms of reference (TOR). The TOR, also commonly called a scope of work, is a key part of the preparation stage. It is effectively a guide to the evaluation describing the objectives, deliverables, methods, activities and organization of the intended evaluation. The more preparation and thought that go into the terms of reference, the more likely the evaluation will be used to shape and inform decision making in the future.

It is effectively a guide to the evaluation describing the objectives, deliverables, methods, activities and organization of the intended evaluation.

The TOR is not only the "evaluation guide," it is also commonly used as the basis for the contract with evaluators and as part of the recruiting materials for prospective consultants.

When should the terms of reference be developed?

Evaluation management should begin four to six months before the evaluation is to begin. The longer the timeframe you have to develop the terms of reference, the less intense the work will be, which will make it easier to weave the tasks related to the evaluation into existing workloads. In addition, the more opportunity there is to reflect and make the most useful choices, the more time there will be to recruit the most qualified evaluation team.

Experienced evaluators are generally not available to work on an evaluation with less than two-three months notice. Therefore, the more time they have between when they receive notification about the evaluation and when the evaluation is expected to start, the greater the chance that the evaluator of choice will be available during the desired timeframe.

Who is involved in developing the terms of reference?

The key actors in the Evaluation Preparation decisions should meet to review their decisions to ensure that project needs are met. This generally includes the project team and the DM&E technical assistant. This is particularly helpful if new staff have joined the team after the design stage as they are able to add their input. This same group does not need to jointly manage the entire evaluation; however, they should be involved, at a minimum, in determining the Lines of Inquiry and in discussing the Evaluation Methodology. Their contribution can continue beyond this, although it may not be feasible in terms of balancing workloads.

Evaluation management should begin four to six months before the evaluation is to begin.

The key actors in the Evaluation Preparation decisions should meet to review their decisions to ensure that project needs are met.

At this point, an Evaluation Manager should be appointed to oversee and guide the evaluation from this time forward. Ideally, this person should not be part of the project to be evaluated. This is an important role that creates a champion for the process and distinguishes between those being evaluated and those who are responsible for the quality of the evaluation. The establishment of this position should in no way exclude the project team from the process; rather, provide a buffer between them and the evaluators.

Some of the decisions necessary to complete the TOR can often be better made in conjunction with input from the evaluation team. In this case, the organization might create a "draft" TOR for recruitment of evaluators who would then help complete the TOR once they have been hired. This is most commonly done for the evaluation methodology decisions. For instance, deciding how many sites of a multi-site project in Angola need to be visited may be made easier with the evaluators' input.

It is important that the terms of reference be finalized and agreed upon by all parties before the evaluation team begins collecting data. Additional changes may be required as the process unfolds, which should be made in writing with the agreement of all parties involved.

It is important that the terms of reference be finalized and agreed upon by all parties before the evaluation team begins collecting data.

What do the terms of reference contain?

The terms of reference can be organized in many different ways and the sections vary by agency. The most common sections are discussed below. What gets fed into each section of a TOR should be tailored to each new evaluation.

Overview

This summarizes the key elements of the evaluation including:

- Evaluation type
- Dates
- Duration
- Intervention
- Country
- Request for applicants

Background

INTERVENTION SUMMARY: This is a maximum of two pages in length and can be supplemented by the completed management tool (logical framework or results framework). The organization can attach it as an appendix to this section or make it available upon request. This section covers key information on the project such as:

- Project goals
- Project objectives
- Current activities
- Location of activities
- Target audience
- Size or scale of project
- Duration of project

ORGANIZATION OVERVIEW: This offers key information about the organization including:

- Organizational mission
- Years in existence
- Size in staff, country offices or budget

Terms of Reference Standard Sections

Overview
Background
  Intervention Summary
  Organization Overview
The Evaluation
  Evaluation Goal
  Evaluation Objectives
  Lines of Inquiry
  Audience
  Evaluation Methods
Implementation Information
  Evaluation Manager
  Location
  Deliverables
  Duration and Working Days
  Deadlines
  Logistical Support
The Evaluation Team
  Role of Evaluators
  Evaluator's Responsibilities
  Evaluator's Qualifications
Application Guidelines
Budget Guidelines
Contact Details

- Length of time in country
- Types of programming offered

The Evaluation

EVALUATION GOAL: This indicates what is ultimately sought from the evaluation. This manual operates from the premise that the evaluation goal is "to improve peacebuilding programming practically and conceptually," though other goals are possible.

EVALUATION OBJECTIVES: This lists the evaluation objectives. These represent what the project team has deemed they want to learn through the evaluation process.

LINES OF INQUIRY: The evaluation objectives set the theme(s) for the evaluation. To ensure that the exploration within that theme delivers useful information to the project team, lines of inquiry are added. Lines of inquiry are a series of questions developed to provide greater direction on what one wants to find out – facts, opinions, experiences, unintended effects, etc. They are often included in the terms of reference in combination with the evaluation objectives. The table on page 142 offers illustrative lines of inquiry for each of the potential evaluation objectives. These should be used to prompt project teams to consider the various issues that an evaluation can explore, though this list should not be considered exhaustive.

Lines of inquiry are a series of questions developed to provide greater direction on what one wants to find out – facts, opinions, experiences, unintended effects, etc.

When project teams generate their lines of inquiry there can be a tendency to list an endless number of questions. To aid in prioritizing this list, especially if it is long, consider how the resulting information will be used and by whom. Questions that do not inform an identifiable decision should be moved to the bottom of the list and, if resources are scarce, should be removed.

Illustrative lines of inquiry

Evaluation Objective	Illustrative Lines of Inquiry
Appropriateness Consideration	• What was the quality of the conflict analysis? • What is the link between the intervention strategy and the conflict analysis? Is it direct? • Does the strategy reflect key leverage points for change? • What is the theory of change? Was it articulated? • Was it appropriate vis-à-vis the context and the intervention strategy? • Are there other strategies/theories of change that could have contributed in a more significant manner?
Strategic Alignment	• How does the intervention contribute to the achievement of the organization's mission? • Does the intervention capitalize on the agency's unique expertise or experience? • Are there other organizations that could do this project better due to their expertise or situation? • Was the implementation reflective of the principles of the organization, for instance gender equality?
Management and Administration	• Was the direction, supervision and support provided to the intervention staff appropriate? • Were all aspects of the intervention well organized? • Were monitoring systems utilized to guide decisions and support reporting? • Were working relationships with partners effective (e.g., good communication, role clarity)? • Were all the activities run that were included in the project work plan? If not, why not?
Cost Accountability	• Were costs projected accurately and tracked regularly? • Were alternatives considered to maximize the use of funds when designing the project? • Did management decisions result in significantly higher costs? • Does the organization try to use economies of scale where possible?

Illustrative lines of inquiry

Evaluation Objective	Illustrative Lines of Inquiry
Process Implementation Appraisal	• Did the development of the process benefit from the lessons from past experience? • Did the staff adhere to 'good practice' standards (where they exist)? • Was the process responsive to the changing context and needs of the stakeholders? • What lessons can be learned from this process? • Did the staff utilize monitoring mechanisms to inform their process? • How was the organizational approach or methodology incorporated into the work?
Output Identification	• What outputs were produced? • Were they of appropriate quality? • Were the outputs in line with the expectations based on the planning documents? If not, why?
Outcome Identification	• What intended outcomes occurred? • Did the outcomes align with the expectations based on the planning documents? If not, why? • What unintended positive and negative changes occurred ? • What part of the project was most important in catalyzing the change? • What was the process (environment, community, personal) that catalyzed the shift? • Was a prior smaller change required to happen first before this outcome could occur?
Impact Assessment	• Is there evidence of the outcomes of the intervention being utilized? • Is there evidence that a transfer of the change occurred from the participants to the wider community? • Were there any unintended negative or positive ramifications occurring due to this project? • What change in the conflict resulted due to the intervention?
Adaptability of Change	• To what degree occurred as a result of the intervention continue? • If the changes were not sustainable, why not? • What could have been done to sustain the results better? • When the conflict shifts to a new phase in the conflict life cycle, will these changes adapt and continue? • Did these changes affect participants' perceptions of what is possible in transforming the conflict?

AUDIENCE: Indicates who the primary audience (users) will be as well as the secondary audience (readers).

EVALUATION METHODS: This section includes the evaluation approach, scope and potentially all or some of the evaluation plan. It also references the existence of monitoring data and the organization's expectations or standards for evaluations.

The approach and the scope of the evaluation should be clearly described. If either requires the project staff to have a role in the evaluation, such as conducting interviews or shadowing the evaluation team, this should be stated. If an evaluation plan exists or there are preferences that will affect the development of the plan such as having a quantitative-qualitative blend of data collection tools, then this should be included. More detailed information on evaluation plans can be found on page 153.

This section should also make reference to the monitoring data that is available to the evaluators. What data has been collected, when, and by whom should be clearly indicated. Finally, if an organization has expectations or standards to which evaluators are to conform, these should be referenced. Expectations range from the requirement to include quotes and stories in the final report to process issues such as when the project team must be included in decision-making. If a formal document exists, attaching them as an appendix to the TOR is a good idea. (To learn more about strategies for avoiding pitfalls, see page 172.)

Expectations range from the requirement to include quotes and stories in the final report to process issues such as when the project team must be included in decision-making

Implementation Information

EVALUATION MANAGER: This details who will manage the evaluation. At a minimum, the person fulfilling this role will be in charge of hiring the evaluators, approving the evaluation plan, ensuring that milestones are met and signing off on the final evaluation.

LOCATION: This section is directly related to the evaluation scope. In effect, it provides the implementation plan of the scope. Indicating the location may be as simple as the physical sites of the project. With a multi-site program the staff team should consider if the evaluators need to access all sites or a representative sample. This is another section that can be finalized with the evaluators' input.

As a rule of thumb, the evaluation report takes about 30% of the total time allocated for the evaluation.

DELIVERABLES: This indicates the products that are required from an evaluation. While there are a variety of deliverables possible, the most common is a report. If a report is desired, the organization should outline the sections it expects to receive. A standard evaluation includes:

- Table of contents
- List of acronyms
- Executive summary of no more than five pages
- Overview of the context
- Organization and program background
- Evaluation methodology
- Evidence-based conclusions: These include the findings (e.g., data) and the analysis (e.g., what they think it means)
- Recommendations (if these were included as part of the evaluation objectives)
- Appendices
    - Appendix A – Terms of reference
    - Appendix B – List of those interviewed (if not confidential)
    - Appendix C – If there are dissenting opinions, they could be included in an appendix

As a rule of thumb, the evaluation report takes about 30% of the total time allocated for the evaluation. For projects operating on tight budgets, one way to stretch that budget is to shrink the expected deliverables. Options include:

- Limited report: A document that offers conclusions and recommendations only, omitting such areas as context and project his tory, methodology, appendices, etc.
- Presentations: Conclusions presented in the form of a presentation with brief handouts to cover the main conclusions
- Informal Format: Use bullet points, rather than narrative style

If a shorter document would increase readership internally, suggest a page limit to the evaluators. Alternatively, one can request a full report plus a summary document which generally averages ten pages in length and provides more detail than an executive summary. The latter option, although very useful since it provides content for a variety of audiences, requires more time from the evaluators and hence has budget implications.

In addition, the report language(s) and the format (electronic, bound hard-copy) in which it is to be submitted should also be specified. If the length of the report is an issue, be sure to indicate what that should be, such as no more than 50 pages for the body of the report, with a five-page executive summary.

Direct quotes or stories can provide insights on the conclusions and bring the generalizations alive. They can also be very useful for other materials such as brochures or the organizational website. If so desired, state specifically that these are expected to be included. This is often found in the Evaluation Expectations document. Note, however, that in conflict situations quoting people by name or using other identifying traits may put them at risk. Furthermore, good evaluation practice demands that the evaluators request permission to reference information disclosed in conversation. Be sure to discuss the desire for quotes with the evaluation team to ensure that this request is feasible within the conflict dynamic.

Finally, indicate who will be responsible for approving the draft report.

Finally, indicate who will be responsible for approving the draft report. This often includes the evaluation manager and project team leader, although such approval may require thematic technical assistance (i.e., review of specific sections or topics by specialists in those subject areas).

DURATION AND WORKING DAYS: The duration of an evaluation considers the period of time in which the evaluation is running, which is often a far longer period than the number of working days. For example, an evaluation which requires 20 working days to complete may start in May with the final report due in the middle of July. Thus, the duration is May to mid-July.

The terms of reference should detail the approximate number of working days needed to complete the evaluation. If international evaluators are being recruited, the working days should be broken down by the number of days they will spend in the country of the evaluation and the number of days they will spend doing work in their home country(ies). These decisions directly affect the budget since in-country days are more expensive. The number of working days depends upon the complexity of the evaluation, the types of information to be collected, the security situation, geography, and the size of the evaluation team.

Sometimes the evaluators need to be in-country to collect data during a specific time period. This could be due to key staff availability, the need to observe an event or project activity, or because of contextual factors such as elections or memorial days that should be avoided. If the in-country dates are fixed or, conversely, if there is flexibility in the schedule, indicate this in the TOR.

One way to estimate the number of working days needed is to allocate days to each task in the evaluation process. A generic task list can be found in the Evaluation Working Days Worksheet below. Since only some of these tasks will be included in every evaluation, this worksheet should be completed near the end of the evaluation planning process.

Evaluation Working Days Worksheet

Tasks	Time Required
Discuss and finalize the terms of reference with project team and/or evaluation manager	
Review organization and project documentation	
Conduct literature review of trends and standards in the conflict transformation field	
Meet with project team to review evaluation process. Discuss the benefits and concerns related to the evaluation.	
Design evaluation plan and discuss with project team	
Design data collection tools	
Test data collection tools	
Travel to and from country	
Travel within country	
Translate data collection tools	
Meet with donors	
Data collection in site one	
Data collection in site two	
Joint evaluation team analysis	
Prepare for project team discussion on draft conclusions	
Draft conclusions presentation and discussion with project team	
Write draft report (rule of thumb for entire report: 30% of overall time)	
Edit and finalize report	
Conduct workshop on evaluation with headquarters staff and/ or key project team staff	
Unexpected time delays (e.g., deterioration of the security situation that delays in-country travel)	
Total time required	

DEADLINES: There are a series of deadlines to set throughout an evaluation including:

- Evaluator recruitment deadline: If recruiting externally, determine when the applications or proposals are due to the evaluation manager.

- Deadline for the proposed evaluation plan.
- The deadline for the draft report.
- The deadline for the final deliverables.

If the evaluation is complex, involves a large evaluation team, or is extended over a long period of time, having a larger number of process deadlines (milestones) is useful for the evaluation manager. These clearly communicate expectations and enable the manager to monitor progress against expectations.

It is smart to set the final report deadline a few weeks ahead of the real deadline since there is a tendency for evaluations to take longer than expected. When setting the deadline, consider proposal deadlines, strategic planning dates, donor reporting cycles and any other process that the evaluation information is expected to feed.

LOGISTICAL SUPPORT: Evaluators rarely operate totally independently of the organization they are evaluating. The organization can save precious time (and therefore money) by providing logistical support to the team. This support might include:

- Ensuring that the team receives key documents in a timely manner
- Organizing travel and accommodation
- Arranging meetings with the project team and key staff
- Identifying qualified translators or local staff to work with the evaluation team
- Providing identity cards where necessary
- Providing mobile phone(s)
- Providing administrative support such as photocopying, fax machines, and office space
- Helping to set up interviews or focus groups, if requested by the evaluation team. (See the discussion of pitfalls, page 172, for more information.)

When an organization does not have the spare capacity to provide logistical support, this should be clearly indicated in the terms of reference so that the evaluation team can plan accordingly.

The Evaluation Team

EVALUATOR'S ROLE: This indicates the role of the evaluation team. (See page 96 in the Evaluation Preparation chapter for further information.)

EVALUATOR'S RESPONSIBILITIES: This outlines the organization's expectations of the evaluation team during each stage in the evaluation process, from developing the evaluation plan to finalizing the report.

EVALUATOR'S QUALIFICATIONS: Many of the key decisions in this area are made in the evaluation preparation stage, such as whether to hire internal or external evaluators, whether the focus will be on local and/or international candidates, the size of the team, and the general types of experience required. The desired and required qualifications must be narrowed down for the terms of reference.

There are a few generic qualifications that should always be considered regardless of the evaluation objectives. These include:

- Evaluation expertise, in terms of both experience and credentials
- Experience working in conflict contexts
- Facilitation skills
- Oral and written communication skills
- Conflict transformation knowledge
- Country experience or, at a minimum, regional experience
- Language capabilities

In addition to the general qualifications that should be considered, there are three considerations that may impact the qualifications best suited for the job: evaluation objectives, gender balance and the conflict context. The following table, Evaluation Objectives & Illustrative Qualifications, offers examples of the different types of experience and knowledge one should consider in relation to each evaluation objective.

This table is not comprehensive nor should it be considered as a list of requirements. In reviewing the table below, the use of the terms "knowledge" and "experience" is intentional. The term "knowledge" implies a minimum understanding of the concepts, while the term "experience"

Evaluation Objectives & Illustrative Qualifications

Evaluation Objectives	Examples of Qualifications
Appropriateness Consideration	• Peacebuilding theories of change knowledge • Theory-based evaluation knowledge • Conflict analysis knowledge
Strategic Alignment	• Knowledge of the organization's history and principles • Knowledge of non-profit performance measurement • Facilitation skills • Conflict transformation knowledge
Management and Administration	• Operations experience • Knowledge of good management practice • Knowledge of non-profit performance measures
Cost Accountability	• Knowledge of good management practice • Basic understanding of accounting systems
Process Implementation Appraisal	• Conflict transformation knowledge • Tool (e.g., dialogue, problem solving workshop) experience • Observation experience
Output, Outcome, Identification	• Tool (e.g., dialogue, problem-solving workshop) knowledge
Impact Identification	• Research skills, ideally quantitative and longitudinal studies • Long term country knowledge
Adaptability of Change	• Country or regional knowledge • Conflict transformation knowledge

It is often difficult to identify evaluators with the full set of required qualifications, particularly when one of the requirements is fluency in a language that is not widely spoken. There is often an inverse relationship between the length of the list of required evaluator qualifications and the length of the list of qualified evaluators. In other words, the more qualifications needed, the smaller the pool of appropriate candidates will be. When faced with this scenario, the organization should reflect carefully on its needs and may wish to reallocate qualifications between the required and desired.

Consider, for example, an evaluation of a peace journalism project in Burundi with a process implementation appraisal objective. Depending on the local situation, it may not require the evaluator to engage extensively with listeners in the community who mainly speak Kirundi. If the majority of the media products are produced in French, the evaluator would not need to speak Kirundi but could conduct the evaluation in French. Speaking the local language would therefore be desired but not required.

Also, consider gender balance on the team, particularly if the organization is committed to gender mainstreaming. Might different genders have differing access to information on the ground, either more limited or more open? Consider, for example, a reconciliation project between two villages in rural Kosovo. Due to the traditional cultural context of the area, male evaluators may not be able to talk alone with the women of the villages. This may cause the women to self-edit their comments in front of men or limit the data collection to the perspectives of the men of the villages only.

Finally, think about the conflict setting when selecting evaluators. Some nationalities or identity groups may have more access or obtain more honest information from the participants of a project or local communities than others. For instance, one would not want to have Palestinian evaluators assessing a project that involved Israeli settlers. In communities traumatized by violence, identity or nationality can be key factors in establishing trust and, therefore, access to useful information.

Think about the conflict setting when selecting evaluators. Some nationalities or identity groups may have more access or obtain more honest information from the participants of a project or local communities than others.

In addition, it is important to remember that the community or participants often see the evaluator of a project as connected to or part of the project being evaluated. As such, if the choice of evaluator implicitly suggests bias or allegiance to a group, this message may also be applied to the project or organization.

See on page 163, Frequently Asked Questions section for further information on selecting qualified evaluators.

Application Guidelines

This section details how interested external evaluators can apply for the evaluation. There are three common ways to recruit external evaluators: sole sourcing, short-lists and open tender. The first, sole sourcing, involves offering the consulting position to one person who is generally well known to the agency. This approach is time efficient, though it requires the organization to know a qualified candidate.

The second option is to create a short-list of candidates and to ask each of them to apply. This request should include the terms of reference, which detail the application expectations and procedures. The short-list can be developed through organizational networks, past consultants or by contacting other agencies for their recommendations.

This approach is less labor intensive than an open tender and more intensive than sole sourcing. It can be useful for those who do not have a candidate in mind and as a way of expanding the potential pool of evaluators for the organization. From the perspective of the short-listed candidates, taking the time to develop a proposal is appealing

if they know that there are a finite number of potential candidates. As such, it increases the chance of quality evaluators taking the time to develop a proposal.

The third approach is open tender, which is similar to the process used to recruit a staff member. The terms of reference are circulated widely and interested candidates are invited to apply. Casting the net wide can be a valuable strategy as it may unearth talent previously unknown. On the other hand, this approach can be time consuming if there is a high volume of interested candidates.

For both the short-list and the open tender approach, the application process can be as simple or complex as necessary. Consider what information you need to select the best candidate and design the process accordingly. The simplest and most time-effective route is to request a CV/resume. While this provides information on a candidate's work experience, it does not give information on the person's approach or aptitude, which is an important limitation.

The rule of thumb is that the more one expects from the evaluators who bid on a project, the fewer applications will be received.

Alternatively, one can request a short – two-page – proposal in addition to the CV/resume. The requested content of the proposal should illustrate the candidate's key skills and experience that are relevant to the evaluation. For instance, the candidate could contribute a draft evaluation plan or indicate the challenges they expect to face and how they would overcome them or they could be asked to articulate the principles that guide their work. Expectations should be kept fairly low if requesting an evaluation plan since it will only be a "best-guess" from the candidates, who are operating on the limited information available in the terms of reference.

Finally, one can request a complete proposal including evaluation plan, budgets, CV/resume and work plan. This can provide valuable information to the selection process, but it will require more time from the selection team. The rule of thumb is that the more one expects from the evaluators who bid on a project, the fewer applications will be received.

For both the short-list and open tender approach, the next step is to identify the two or three best candidates and contact them to assess their competency. For those recruiting from a local pool of candidates, requesting presentations on similar topics may also be an option. More information on assessing competency can be found in the Frequently Asked Questions section on page 163.

Budget Guidelines

It is good practice to include the evaluation's financial parameters in the Budget Guidelines section, since these can provide a sense of the project's size to prospective consultants which, in turn, can enable them

to tailor their application accordingly. For larger budgets, one can include the actual cost figure of the evaluation or provide some guidance on the range of the budget: "The budget for this evaluation is between $70,000 and $85,000 Canadian dollars." Generally, proposals will use the full budget provided. Consequently, if limiting costs is an important criterion in the selection process, it should be clearly indicated.

The project team, the evaluators or a combination can develop the evaluation plan.

If the organization requests budgets to accompany applications, outline what should be included or excluded from the budget along with any other parameters.

Contact Information

If the evaluation manager or project team personnel are willing to field calls from interested individuals, include their contact information in the announcement. Taking inquiries from interested parties can help the evaluation manager gauge both the volume of interest in the project and the expertise available, and the evaluators can become better informed about the proposed evaluation. That said, this can take an inordinate amount of time and, if started, should be offered to all who inquire.

II. THE EVALUATION PLAN

What is an evaluation plan?

An evaluation plan provides a structured layout for designing an evaluation. It facilitates consideration of the key aspects of the evaluation and can be a useful communication tool among the various stakeholders. The evaluation plan also provides a reference point for the evaluation manager and project team to use to contribute to the evaluation design and to monitor the implementation.

Evaluation Plan

Evaluation Objectives	Lines of Inquiry (Indicators, Standards)	Decisions to Inform	Means of Verification (methods)	Data Source & Quantity	Location of Data Collection	Conflict Considerations	Means of Analysis	Time (days)

To put it in context, an evaluation plan is to an evaluation as a work plan is to a project. The evaluation plan should be thought through as much as possible before getting started, yet it should have enough flexibility so that the evaluators can make alterations as they proceed and have the opportunity to identify better means or new opportunities.

Who develops the evaluation plan and when?

The project team, the evaluators or a combination can develop the evaluation plan. If it is to be developed by the project team, that fact should be included in the terms of reference. In such cases, it is advisable to still leave room for the evaluators to suggest improvements, particularly as they gain knowledge about the project and its context. If the project team feels strongly about certain aspects of the plan, such as the location of data collection, for example, be sure to indicate that to the evaluators.

Alternatively, the evaluation plan can be requested as part of the tendering process. The plan can only be as good as the amount of information that the evaluator has, so one should keep expectations reasonable. In this case, it should be expected that the evaluation team will wish to update it as more information becomes available. Finally, the plan may be developed from scratch once the evaluators are contracted.

Whether developed from scratch or as part of the tendering process, the plan should be created in conjunction with the project team whenever feasible. If there is room in the budget, plan development can be highly participatory and can produce a second deliverable of increased organizational capacity. If finances are limited, some strategically timed meetings should be held with the project team to gather their input and concerns before decisions are finalized. Engaging the project team so that they support the methods is a key component of having the evaluation conclusions accepted and applied in the future. See Evaluation Management page 174 for further discussion on this topic.

What does an evaluation plan contain?

While there are a number of different versions, the one depicted in this manual offers a structured approach that includes the core subject areas in most evaluation plans.

EVALUATION OBJECTIVES: These should come directly from the terms of reference. The evaluation team may raise useful questions that will help refine the objectives.

LINES OF INQUIRY, INDICATORS, AND STANDARDS: The evaluation objectives will dictate which of these three are needed. For instance, all results identification objectives must have indicators. If the indicators do not contain targets (e.g., "35% increase in freedom of movement" – with 35% being the target), standards of achievement must be developed for each indicator. Sometimes these

concepts are also blended. For example, lines of inquiry may be sufficient on their own or they may require standards and indicators within each inquiry area.

- **LINES OF INQUIRY**: Lines of inquiry provide the evaluators with greater direction about the information needed by the project team. Whether indicators or standards are used within each inquiry area depends on the evaluation objectives. See the table on page 142 for an illustrative listing of lines of inquiry.

- **INDICATORS**: Ideally, the indicators are drawn from the project design documents, though the evaluation team may wish to add to or improve them. Adding new indicators is fine as long as it is permitted by the grant agreement. Some grants require the implementing partner to gather information on specific indicators as part of a larger cross-agency monitoring system. Another consideration is whether or not the project monitoring system has been operating effectively. If so, data related to existing indicators will already have been collected, so by changing or deleting them, that information may no longer be useable.

- **STANDARDS**: There are two forms of standards, process standards and standards of achievement, and both are used as a comparison against which the evaluator can draw conclusions. Where the evaluation objectives include process implementation appraisal, management and administration, or cost accountability, process standards must be established. These standards may comprise techniques, steps, procedures, principles, or some combination of them all.

  In fields like public health, there are international standards of practice that can be utilized. Peacebuilding practice does not yet have internationally accepted standards of practice against which quality may be assessed. Consequently, each organization needs to spend time reflecting on what constitutes high quality process to enable the evaluation to provide useful information. Experienced evaluators may be able to facilitate this discussion.

  Standards of achievement are also used to compare findings against. When one of the evaluation objectives is results identification, the evaluator can draw on the targets within the indicator as a point of comparison. Consider this example: "An expected 35% increase in young Nepalese men from village Q who openly travel through the "other's" community at night." If the evaluator finds that there has been a 50% increase in such travel, the 15% difference is obviously far above the intended standard or target set by the program. The key point is that there is something to compare the evaluator's findings against in order to draw a conclusion.

Peacebuilding practice does not yet have internationally accepted standards of practice against which quality may be assessed.

If there is no target in the indicator, time should be spent at this stage determining what would constitute success. Although largely driven by the project team, stakeholders' views are extremely valuable. The evaluator's experience can also be quite informative. Of course, purposely setting a low standard is not only unethical, it is also unlikely to work. An experienced evaluator will have had other scenarios to compare the current one with and will make note in the evaluation of the below-average target.

If there is no target in the indicator, time should be spent at this stage determining what would constitute success.

DECISIONS TO INFORM: In order to achieve the evaluation goal of improving peacebuilding programs practically and conceptually, the evaluation plan must clearly relate to the decisions or learning to be sought from the evaluation.

MEANS OF VERIFICATION (MOV): The MOV, or data collection method, is the way in which data will be collected. Different methods may be selected for each evaluation objective or line of inquiry/standard/indicator or one method may be appropriate to gather information against a number of them. Since each method must be developed and then tested, the more methods utilized the more time required for the evaluation. One can also use multiple methods to triangulate the information received as a way of verifying its accuracy.

The MOV, or data collection method, is the way in which data will be collected.

All social science data collection methods, such as interviews, questionnaires, document review, or focus groups are potential means of verification, as are participatory methods such as mapping or drawing. Further information may be found in the Methods chapter on page 204.

In the last several years, there has been increasing attention on developing conflict transformation-specific methods that will meet the special needs of this work. Many of these are peacebuilding-specific modifications of a social science methodology such as questionnaires, while others establish their own processes. As of 2005, none of these newly devised tools have taken precedence, nor have many been thoroughly tested. There is every indication that this area will continue to receive increasing amounts of attention, which will only improve and expand the options. Since this manual is focused on designing for results rather than on conducting the evaluations themselves, listing the various new tools is beyond its scope. Further information on peacebuilding tools may be found in the Methods chapter on page 204.

In the last several years, there has been increasing attention on developing conflict transformation-specific methods that will meet the special needs of this work.

The data source is where the information will be accessed

DATA SOURCE & QUANTITY: Closely tied to the MOV section, the data source is where the information will be accessed, whether from participants in the project, media professionals, judges or schoolteachers. Again, the data source can be different for each objective or line of inquiry/standard/indicator or these can overlap. The key question is how the evaluator will best access the information. This section should be as specific as possible. Indicating "women", for example, as the data source is too general and will hinder the subsequent decisions in the evaluation plan.

This column should also give an estimate of the number of responses needed from that data source. For instance, 90-120 (quantity) questionnaires (method) from NGO staff members whose agencies currently work on the flashpoints in Belfast and Derry/Londonderry, Northern Ireland (data source). The quantity needed will depend upon the scope of the evaluation. If the evaluation is national in coverage or requires generalized conclusions, the quantity necessary will be higher.

LOCATION OF DATA COLLECTION: At a general level, where data is collected will already be detailed in the terms of reference. The parameters, however, are often based on areas larger than the evaluator may be able to cover, which means that a second set of decisions must be made to more precisely pinpoint where the data collection will take place.

Is this realistic within the conflict context in terms of resources, opportunities and constraints?

CONFLICT CONSIDERATIONS: This section is intended to systematically insert the conflict into each layer of the evaluation planning. Fundamentally, one needs to ask throughout the development of the evaluation plan: Is this realistic within the conflict context in terms of resources, opportunities and constraints? The project team is particularly valuable in this regard since they not only know the answers but often are better equipped to identify the salient questions. As decisions are made within the evaluation plan, check them against these additional questions:

- What is the security situation? Will it restrict travel? Will it restrict access to people? Can some people move with more security then others?
- Can the evaluation team physically gain entry into the target community?
- Do special measures need to be taken in working with a community such as a highly traumatized group?
- What steps need to be taken to guarantee the safety of those involved in the evaluation?
- What are the prohibited or taboo questions?
- What are acceptable and unacceptable words? What language is considered politically biased?
- In which language will the evaluation be conducted? Languages can have political implications and should be considered carefully.
- What implications does the identity of the evaluator have within a community?

- Can the evaluation process include basic conflict transformation principles such as, "conflict is not always negative" or "generating a best alternative to a negotiated agreement (BATNA)" in order to support the goals of the intervention?

- What impact on conflict will the evaluation actions have?

- What physical infrastructure exists such as postal systems, telephone lines and accessibility of roads?

- What are the literacy levels of the target groups?

- Is this an appropriate method for the culture from which I am gathering data?

- What methods are considered valid in the culture of the evaluation audience?

- What are the cultural expressions (whether physical or verbal) of discomfort or unease?

MEANS OF ANALYSIS: How will the findings be analyzed? Consideration of the means of analysis is an important step yet it is one that is often forgotten when doing qualitative research. This is especially important to consider when the members of an evaluation team, who all bring specific skills and expertise to the effort, have different analytical approaches. Since the analysis approach selected affects the conclusions generated, a team member with a particular background may recommend an analysis method that works best for her/his substance area not knowing that use of that approach will omit important information from other areas. Particular thought should therefore be given to how differences of opinion about the analysis will be discovered and resolved.

TIME (DAYS): This column outlines the estimated time that will be needed to accomplish each task. It is a valuable reality check to ensure that the evaluation plan stays within the projected time period in the budget.

How do you develop the evaluation plan?

Creating an evaluation plan requires a blend of creativity and practicality – creativity in determining the best way to access the right information and practicality in terms of operating effectively within the constraints of time, finances, skills, and context. Generally, finding this balance will require some flexibility throughout the evaluation, but it does not eliminate the need for planning.

Note that the degree of rigor (exactness and complexity) within the research components of the evaluation plan is not that of a doctoral

student. Yet good research practice must be applied so that the conclusions are credible. One standard to apply is: Will this plan provide the organization with sound information that it can use to base decisions upon with confidence?

Evaluation Plan Example

The evaluation plan is for a project located in Bukavu, Democratic Republic of the Congo (DRC) that has been in existence for one year. The project consists of four radio programs whose goal is to help strengthen the people's commitment to the Intercongolese Dialogue. Two of the four activities have been selected for inclusion in the evaluation. The evaluation plan, developed by the evaluator, only shows one aspect of the evaluation and is intended to be illustrative rather than comprehensive.

Will this plan provide the organization with sound information that it can use to base decisions upon with confidence?

PROJECT LOGIC:

Goal: Strengthen the people's commitment to the Intercongolese Dialogue

Objective One: Increase people's knowledge of the Intercongolese Dialogue

Objective Two: Increase people's understanding of the relevancy of the agreements to everyday life

Objective Three: Increase people's participation in the dialogue

Activities: Four radio programs: one soap opera and three talk shows of differing formats and themes

Definition: Intercongolese Dialogue means the agreements and dialogue process

One of the evaluation objectives in the terms of reference is output identification because the project team wants to ensure that the activities are producing deliverables in as intended. In addition, since the project is only a year old, outcomes may exist, but it also may be too early to identify strong evidence of them. Consequently, collecting solid output information is helpful in terms of ensuring that the team is building toward what they hope to accomplish.

The line of inquiry associated with the objective is, "What outputs have been produced by this program over the past year?" More lines of inquiry could be included if the project team needed other information to help them make decisions. For instance, a line of inquiry that looked at the quality of the outputs could be added.

Evaluation Plan Example

Evaluation Plan							
Evaluation Objectives	**Lines of Inquiry (Indicators, Standards)**	**MOV (methods)**	**Data Source & Quantity**	**Location of Data Collection**	**Conflict Considerations**	**Means of Analysis**	**Time**
Output Identification	*Line of inquiry:* What outputs have been produced from this program over the past year?						
	80% of media staff able to independently produce a radio show	In-person interview	Head of radio production Country director	Bukavu NGO office		Review of interview notes	.25 day
	280 minutes/month of radio program Y aired during prime time for the target audience	Document review Radio listening	Weekly production log (one/ month) One station/ radio program	Bukavu NGO office		Collate the consistency of broadcast times and radio minutes Verify broadcasting by tuning in at the right time for each show	.5 day
	50% of radio station listeners repeatedly listen to radio program	In-person interviews	Radio manager(s) from each of the 5 stations	Throughout Bukavu city	Female evaluator should not travel alone Is the population settled enough for these numbers to be gathered?		1.25 days

The indicators were found in the program design tool which was a logical framework. In this case, the indicators apply to both activities included in this evaluation. The original indicators were:

- Number of individuals able to independently produce a radio show
- Number of minutes of radio programs aired/month during prime time for the target audience
- Number of listeners/radio program

If the activities were different, such as a youth camp and radio programming, the output indicators would need to be unique to each of those activities. Minor changes were made to the first two indicators to make them more specific and to avoid any misinterpretation. The third had to be altered to meet the realities on the ground, although its essence was kept.

None of the indicators contained targets, so the evaluation team established the standard of achievement at this stage. This process included individual discussions with the project team, emails sent to headquarters support staff, former staff and to the regional media manager. They were asked to reflect back on when the project was designed and their expectations at that point. Since it is difficult to separate what one knows now from what was known when the project began, all of these responses were brought together and blended with the expertise on the evaluation team. The results were then brought back to the current project team to ensure that they accepted them as an appropriate standard.

The data on the first indicator, "80% of media staff able to produce a radio show," would be found through two in-person interviews (method), one with the head of radio production and one with the country director (data source and quantity). Since this is a new office, open only for one year, the team is small enough for these the head of radio production and the country director to work directly with each staff person. Therefore, the evaluators felt confident that they were an adequate source of information; however, it would be necessary to verify this in order to rule out any potential staff politics at play. To this end, the evaluators reviewed the radio production log to determine the number of shows produced by each staff person, frequency of production, and variety of working partnerships. The latter was to ensure that a team had not formed where one person actually produced the show without much assistance from the other. Note that, if there was a quality component to this indicator (e.g., number of individuals able to independently produce high-quality radio shows), standards would need to be the established to indicate what constitutes "high-quality," and the evaluators would then need to assess the staff against these standards.

The two interviews can be conducted in the Bukavu office (location of data collection), though they should be held behind closed doors since they may involve anonymous comments about staff competency. The means of analysis in this case is very simple and can be a comparison of the interview notes. This will answer the question, "Do we need to provide additional staff training in radio production?"

The evaluation team questioned the second indicator, "280 minutes/month of radio program 'Y' aired/month during primetime for the target audience," in terms of whether it was a result of or simply a part of project activities. After discussion with the project team, it was accepted as one of the indicators to be used because the project produces shows and does not have its own radio station. As a result, part of the achievement in this case is convincing radio stations to air their broadcasts during the target audience's prime listening time. Of course, if the organization pays the station to play its programming, this would not be an indicator since would simply be part of the activities.

The data on this indicator can be gathered by reviewing (method) twelve weekly production logs (data source and quantity) – one per month. One does not need to review every week, unless it is evident from the sample selected that there is great variance in the information. Since the monitoring needs informed the creation of the production logs, data on when the radio show was broadcast each week is recorded in the log. To verify that this is accurate, the evaluator could tune in (method) to each of the two programs (data source and quantity) at their scheduled broadcast times. This can all be accomplished from the Bukavu office of the NGO (location of data collection). The production log information would be collated to determine the consistency of the broadcast times and the number of minutes added up for an approximate total (means of analysis).

This information will indicate to the project team whether or not they need to do additional work with the radio stations in order to get the correct time slot for the target audience. It will also contribute valuable information to the NGO's understanding of the intensity and frequency of messages required to initiate change.

The third indicator is "50% of radio station listeners repeatedly listen to radio program Y." Although seemingly straightforward, it is very difficult to gather accurate data for this indicator. If the evaluation budget and time were sufficient, one could conceivably gather this data in a normal context. However, Bukavu has limited road infrastructure, a highly insecure environment outside of the city itself, and an extremely mobile population (conflict considerations), all of which would make it difficult for the team to generate reliable conclusions. With sufficient time and money, these challenges could be surmounted to some degree; however, this budget did not offer that opportunity.

The fall-back alternative is to conduct interviews (method) with the radio managers (data source) at all five stations (quantity) to seek their assessments of the percentage of listener market share each of the stations has. Since there are no land lines in Bukavu and few cell phones, these need to be in-person interviews at the stations (location). In addition, the approach to these interviews should be flexible because there may be multiple managers or other knowledgeable personnel at each station, which may require small group interviews to be conducted instead. Due to the high levels of violence against women, the female evaluator should not travel alone to these stations (conflict consideration).

Although it is almost a certainty that the stations will not have statistically generated ratings, they will have an informal sense of their stations' audience and the popularity of each show. This data should be supplemented by including questions in the other aspects of the evaluation to triangulate the estimates. The means of analysis will be to compare the radio station estimates of listenership with the other data gathered. The evaluation team should indicate these trade-offs to the project team so that they understand the limitations of the information. All of this information will feed into the organization's decision regarding whether or not to continue airing each of the radio programs.

III. FREQUENTLY ASKED QUESTIONS ON WORKING WITH EXTERNAL EVALUATORS

Once the decision has been made to use an external evaluator (see page 126 for more information on this decision), a variety of questions are commonly raised. This section provides some tips and tactics for answering these questions.

1. When should I start recruiting to get a quality evaluator?
2. Where do I find evaluators?
3. How do I appraise the competence of external evaluators?
4. Is there anything different in an evaluator's contract?
5. What do I need to do before the evaluation begins?
6. How do I know if the evaluation plan is a good one?
7. The security situation has worsened and it is not safe for the evaluation team to travel to some of the worksites. Should we cancel the evaluation?
8. Who should manage the evaluators?
9. How much management is needed for external evaluators?
10. Can I participate in the evaluation of my project?

11. What should I do if I don't agree with the draft report?
12. What do we do with the evaluator's primary data, such as interview notes, after the evaluation is completed?

1) When should I start recruiting to get a quality evaluator?

Whichever approach is selected for recruitment, sole-source, short-list or open tender, professional evaluators with track records of quality performance generally have booked schedules. Therefore, to obtain an experienced professional, it is recommended that the recruitment take place as far in advance of the evaluation as possible. In the ideal scenario, development of the TOR should begin six months before the time the team is expected to arrive on the ground. As a rule of thumb, there should be two months between finalizing the contract with the evaluator and the on-the-ground dates of the evaluation.

...professional evaluators with track records of quality performance generally have booked schedules.

Consider, as an example, an evaluation that is to occur in June. The TOR would be developed in January/February, the recruitment process would occur in March, and the contract finalized no later then early April. In May, the preparation should start so that the team can be on the ground in June. Shorter timelines are possible, but the organization will pay the cost in terms of both the intensity of work required to plan the evaluation, which must be squeezed in among regular job responsibilities, and a likely decrease in the quality of evaluators available.

2) Where do I find evaluators?

There are a number of ways to connect with prospective evaluators including evaluation associations, M&E websites, universities or conflict transformation networks.

Many countries have national evaluation associations that offer listservs, job postings on their websites, or on-line databases to connect with their membership. There are also a few M&E websites that post opportunities for evaluators. Check Appendix B on page 228 for a listing of these websites and a selection of national evaluation associations.

To facilitate the sole-source and short list approaches, it is recommended that organizations maintain a database of the evaluators it has used. Include feedback from the project team and evaluation manager on the professionalism, quality and performance of the evaluator to inform future decisions. Names of individuals who have not performed to a reasonable standard should also be recorded to prevent the organization from making the same hiring mistake in the future.

3) How do I appraise the competence of external evaluators?

There are many ways to appraise an evaluator's competence. Some of the more common steps are included here. The first step is to compare the person's application against the qualifications listed in the terms of reference. Be sure that she/he has substantive and solid experience in the required areas as evidenced by the CV/resume. A common mistake is to rely on the cover letter alone, which may provide a different picture from the CV/resume itself.

A common mistake is to rely on the cover letter alone, which may provide a different picture from the CV/resume itself.

The next step is usually an interview with the evaluator or the evaluation team. In this discussion, seek to determine the depth of the person's expertise on evaluation as well as her/his core substantive areas of expertise (such as child soldiers or ADR processes). Find out if the evaluation team wrote the proposal or application; if not, be sure that the proposed evaluators are as qualified as the application writer. In addition, keep in mind that the existence of many completed evaluation assignments on a CV/resume may not be an indicator of quality. As such, it is a good idea to request a sample evaluation to review, though this may be more difficult than expected because the majority of evaluations are the property of the organization evaluated. When reviewing a previous evaluation, check to see if it has the elements of quality that the organization has determined it needs for its proposed evaluation.

When using an open tender or short-list approach, it is always a useful final step to check references. In these conversations, be sure to ask about:

- Timeliness: Did they meet the milestones and deliver the product on time?
- Responsiveness: Were they flexible enough to deal with unexpected challenges or delays? Did they respond to the project team in an appropriate manner?
- Relevance: Did they follow the terms of reference?
- Professionalism: What was their work style, communication ability, or degree of cultural sensitivity?
- Evaluator Good Practice: Did they engage openly with the project team? Did they proactively explain their decisions based on good practice in evaluation?
- Evidence-based Conclusions: Were the report's conclusions evidence-based or mainly conjecture?

- Conflict Sensitivity: Did they display a sound understanding of the implications of doing evaluations in conflict areas?

A common mistake is to simply ask the referee if she/he was happy with the evaluation. This assumes both a level of knowledge about good practice in evaluation and that the two organizations have similar expectations of quality. In some cases, the referee may be happy with the product simply because it reflected well on her/him rather than because the quality of the evaluation was high.

4) Is there anything different in an evaluator's contract?

The majority of a contract for external evaluators is the same as for any consultant. Remember to include the terms of reference as an appendix and refer to it in the body of the contract. If the organization has developed standards or expectations for what constitutes a quality evaluation, these should also be included either within the contract itself or as an appendix.

A section should also be included that indicates how the evaluator can use the information after the project is completed. It is most often referred to as "confidentiality" and it should state precisely who owns the data and the final report as well as how these may be used in the future. For instance, if the organization owns the final report from an evaluation, the evaluator would need to request permission to use it as a sample in an application for another assignment. The same holds true for posting a report or information derived from the evaluation on a website, quoting it in other reports or making public statements about the conclusions.

A common mistake is to simply ask the referee if she/he was happy with the evaluation.

5) What do I need to do before the evaluation begins?

A number of actions must occur between the time that the evaluation team is hired and the start of the evaluation. The first step is to gather all relevant documentation for the evaluation team. This almost always includes the funding proposal(s), donor reports, baseline and/or previous evaluations, program background information, organizational history and monitoring data. The challenge is to provide critical documentation that will inform the team but which will not overwhelm them. Be sure to create a listing of key staff, titles and contact details within the project if one does not already exist.

The evaluation will proceed much more smoothly if it has been discussed with the project staff before the evaluator arrives in their office. Transparency and communication around the motivations for the evaluation, the role of the evaluator, and the potential uses of the final report can allay

natural fears that may arise. These discussions should encourage staff to be open and honest with the evaluator. In this process, individuals who oppose the evaluation may appear, which may require that individual or more focused discussions be had with them prior to the evaluation to allay their fears.

It may be useful to think about which key stakeholders the evaluators should meet beyond the project staff. The list of these stakeholders should then be reviewed with a conflict lens. For instance, will involvement in the evaluation compromise the position of any of these people? Could their participation place them in danger in their community? Are any of the stakeholders so badly traumatized that the evaluators should be careful not to negatively affect these individuals? Is there anyone who may not be willing to meet with strangers (i.e., the evaluators) or trust them enough to speak honestly and would therefore need a personal introduction from a project staff person? Having this list available with any conflict considerations will be very helpful to the evaluators.

Transparency and communication around the motivations for the evaluation, the role of the evaluator, and the potential uses of the final report can allay natural fears that may arise.

Finally, pending what logistical support was promised in the terms of reference some logistics may need to be addressed such as the purchasing of airplane tickets.

6) How do I know if the evaluation plan is a good one?

So, you have left the methodology open for the evaluators to construct and they have submitted an evaluation plan for approval. Here are some of the questions to ask:

- Has each of the evaluation objectives been included?
- Are the lines of inquiry supplemented where needed by standards or indicators?
- If new indicators were added, how are they better? Are they "good" ones?
- Are all the necessary locations covered?
- Are the research methods appropriate for the context? For instance, is a questionnaire suggested for a location that has a low literacy rate? Are focus groups planned in areas where women cannot attend?
- Do the research methods appear to be the best ones for obtaining the data? For example, focus groups are not appropriate if community members tend not to speak openly and honestly in public.

- Is there a quantitative-qualitative blend in the research methods?
- Do the methods triangulate?
- Are there any concerns with accessibility to the data source groups?
- Does the amount of time allotted for the evaluation seem feasible? Are there more days planned than are available for the evaluation? Has the time required for local travel been taken into account? Does the plan account for the extra time that translation requires?
- Is existing monitoring data being used?
- Have they ensured that different sources or different processes have verified the monitoring data?
- Are the conflict considerations accurate, necessary and comprehensive?
- Are the means of analysis benefiting from the full team's expertise?

If the organization has M&E technical assistance available from staff members who are not already involved in the evaluation, this is certainly the time to request their input. If there is no in-house expertise, consider hiring an M&E advisor who can be consulted at key steps throughout the evaluation process. To save time, the same advisor can be used for all evaluations.

7) The security situation has worsened and it is not safe for the evaluation team to travel to some of our worksites. Should we cancel the evaluation?

The answer to this question is, "It depends." When the safety of the team is at stake, one should always consider postponing to a later date. Under circumstances of questionable safety, there are at least five other options to consider: changing the methods, excluding locations, reframing who travels, considering proxies and using secondary sources.

First, are there alternatives to the planned data collection methods that do not require travel to the dangerous area? Consider using phone interviews or a staff person who is already in the area as ways to conduct the conversation and provide notes to the evaluation team. If a case study is being written, key actors can be asked to chronicle their experiences if this would not put them in danger.

Second, can other worksites be used without compromising the results to a great degree? For instance, a nationwide survey on media attitudes in Burundi excluded from the sample the three provinces that were inaccessible because of poor security conditions. Though such exclusion might not be the ideal, it is still a credible option as long as the limitations in the data collection process are clearly communicated in the final report and the conclusions qualified accordingly.

Third, can the people who are to be interviewed travel safely to the evaluation team? In certain cases, movement is possible for one community but not the other.

When the safety of the team is at stake, one should always consider postponing to a later date.

Fourth, are there proxies available who can speak to the evaluator(s) on behalf of the target population? Talking to parents of child soldiers who are fighting in the bush for their perception on the child's attitudes and experience is an example. Clearly, this is not a perfect strategy; however, if one has no way of accessing child soldiers, it may provide insights and information that, when triangulated with other information, becomes useful in decision making.

Fifth, shift the data collection to existing sources of information, such as other NGO studies or internal reports, academic research papers, evaluations for other projects or other organizations, newspaper articles, etc. The exclusive use of secondary sources is a blunt instrument, but if enough of it can be collected, a useful picture can emerge on issues addressed by the evaluation. Which of these strategies makes sense will depend on the evaluation objectives and intended use.

8) Who should manage the evaluators?

Appointing an evaluation manager who is not on the project team of the project to be evaluated is the most effective strategy. When done well, the separation of evaluation management from the actual evaluation creates a "political firewall" between the evaluators and the program, which keeps the evaluation free from the control or undue influence of those responsible for the project. Moreover, this structure enables an internally driven evaluation to be as close to an independent evaluation as possible.

Appointing an evaluation manager who is not on the project team of the project to be evaluated is the most effective strategy.

The evaluation manager is appointed at the beginning and is named in the contract as the responsible party. She/he works with the program team on all key evaluation preparation decisions and on implementation and use (stages one, two, and three). The person in this position facilitates development of the terms of reference, runs the recruitment process, manages feedback on the draft report, signs off on the completion of the final report and authorizes final payment. Good communication by the manager with the program team is key to producing a useful document.

The manager needs to keep squarely in mind that she/he is managing the administration and not the substantive content. The nature of the conclusions is outside the remit of the manager, unless the conclusion does not seem evidence-based or is missing key variables. In that case, it is appropriate for the manager to request an alteration to the report in the form of either additional evidence to support the conclusion or the incorporation of the missing variables, provided that it had been verified.

In a mixed evaluation team, with internal staff and an external consultant, it is helpful if the manager is not the internal staff representative. Separation of those roles will decrease role confusion and possible tension. Where management of an evaluation becomes more difficult is when the individual managing the evaluation is also part of the project being evaluated (see the ethics chapter page 188).

9) How much management is needed for external evaluators?

External evaluators should be managed in much the same way as any external consultant under contract. Ensuring that the team is using its resources wisely, staying aligned with the terms of reference and meeting deadlines should all be managed by the organization.

The management role is a very active one from the preparation stage through the data collection stage. Prior to the point when evaluators arrive on site, the manager should carefully review the draft evaluation plan. Although communication will be ongoing while the team is conducting the evaluation, the manager should hold a final discussion with the evaluators before they leave the field location to ensure that all of the necessary data has been collected. This discussion should also identify for the project team what the next steps will be, including the submission of draft conclusions. For a complicated or lengthy evaluation, periodic milestones should be agreed upon and the manager should regularly check progress against these milestones. The manager continues to ensure that deadlines are met and report specifications are maintained after this point, but the nature of the report's conclusions is outside of the management mandate.

The manager needs to keep squarely in mind that she/he is managing the administration and not the substantive content.

10) Can I participate in the evaluation of my project?

This decision should be made in the evaluation preparation (stage one) and all project staff members should be advised about their respective roles. If it is a self-evaluation, a participatory evaluation, or you are a member of a mixed evaluation team, then you will likely be very involved in the evaluation process. Alternatively, if the evaluation has

been structured to include a capacity strengthening element (similar to a participatory evaluation), there will be structured ways for you to participate.

As a member of the project team, you may also be invited to meetings with the evaluation team. This would keep you informed of the progress and decisions being made. Finally, the evaluation team should be required to provide feedback to the staff before they leave the field. This feedback can range from initial impressions to sound analytical conclusions, depending on whether the in-country team has had time to do the analysis. This is an important step since it provides the staff with an opportunity to question, understand, and clarify different conclusions.

Direct requests to participate from the project staff to the evaluators can be awkward and perceived to be power-politics at work.

Beyond these relatively structured forms of participation, inclusion in the evaluation process should be left to the discretion of the evaluation manager, who will consult with the evaluation team. Direct requests to participate from the project staff to the evaluators can be awkward and perceived to be power-politics at work. Be respectful of the evaluator's role and the need for the conclusions to be sound and the evaluation credible.

11) What should I do if I don't agree with the draft report?

The principle that cannot be violated with an evaluation is that the conclusions must be based on the evidence.

The principle that cannot be violated with an evaluation is that the conclusions must be based on the evidence. These are not the views and opinions of the project staff that have been foisted onto the evaluation team. Nor should evaluators be drawing conclusions that are not based on the data.

That said, there will be instances where the project team does not agree with the evaluation conclusions. If there is disagreement because the conclusions are not supported by evidence, a request can be made that the evidence be included or the conclusion omitted. Of course, this assumes that the evaluation terms of reference included evidence-based results. If it appears that the evaluator was not aware of critical information when she/he analyzed the information, it is appropriate to inform the evaluation team of these extra variables and how you believe it impacts the conclusions.

If the disagreement is with the conclusion itself and the evaluator has heard the concerns and decided not to alter the conclusion, one alternative is to include the organization's view as part of the evaluation document. This can be done in the text, as a footnote, or as an appendix to create a document that shows both perspectives. Including the differing views is extremely important because the next evaluation will likely ask what happened to the conclusions in the previous evaluation.

Editorial control of the final report is often indicated in the terms of reference. Generally speaking, the last word on the final evaluation report should be left to the evaluators. If another version of editorial control is selected – though this is not recommended – it should be specified in the terms of reference.

12) What do we do with the evaluator's primary data, such as interview notes, after the evaluation is completed?

Primary data, such as interview notes, should be destroyed if the evaluator promised confidentiality to the sources in the process. This is a particular necessity in situations of conflict where the opinions and stories collected may be about sensitive issues and, at times, even illegal ones.

Generally speaking, the last word on the final evaluation report should be left to the evaluators.

Hard copies of surveys and questionnaires should be stored with the analysis and conclusions so that future evaluators can use the raw data to verify the conclusions of the first evaluation or to contribute to a different study to save resources. If the data has been entered electronically into a software program, this too should be saved, again to eliminate the need to re-enter it for future uses. If the surveys or questionnaires include the responder's name, the organization must take extra precautions to store that information in a manner that protects the responder.

IV. STRATEGIES FOR OVERCOMING COMMON EVALUATION PITFALLS

There are many ways in which evaluations can falter along the way that can hinder the quality, usefulness or application of the evaluation conclusions. With some preplanning and creative thinking, these common pitfalls can be avoided.

"When the eyes are bigger than the plate"

It is common for evaluation managers and project teams to design an evaluation that is far too large for the proposed budget. Creating too many evaluation objectives or having too broad a scope for what is feasible are the most common ways for this to happen. To avoid this problem, be sure to work through the key evaluation preparation decisions

in the project design process, as outlined in the Evaluation Preparation Decision flowchart on page 99. This will enable the project team to create a sound budget estimate as part of the project proposal.

If the evaluation appears to be too large for the budget and the team is having difficulty scaling back, draft the terms of reference in a flexible way until the evaluators have been contracted. Then, during the first meeting between the evaluators and the project team, they can work collaboratively to finalize the terms of reference with an eye towards feasibility.

"Don't shoot the messenger"

There are two common responses to negative conclusions from an evaluation: discount the methods or criticize the evaluator.

One of the many benefits of evaluation is that it helps to identify issues and information for the project team that were previously unknown. In this process, both positive and negative points can be unearthed. Sometimes the negative findings can be quite challenging and even threatening for a project team to handle and there can be a tendency to resist the resulting conclusions and recommendations.

There are two common responses to negative conclusions from an evaluation: discount the methods or criticize the evaluator. In the first of these, the validity of the methods used in the evaluation will be questioned. Statements like, "Of course that is what you found; you talked to the wrong people!" or "I'm not surprised they didn't tell you otherwise since you spoke to them in a group and they would never contradict the norm in that setting," depict this well. A strategy that may help prevent such reactions is to engage the project team in the methods discussion. This discussion can be as participatory as is feasible and it should be made clear that this is the appropriate time to challenge the methods. At a minimum, a discussion based on the evaluation plan should be held with the full project team with the explicit purpose of getting the team to reach consensus on the best strategy.

The second common response to negative conclusions is to challenge the evaluators' credibility. Statements like, "Well, this field is not their specialty, so they just don't get it" or "I didn't like him from the beginning," are often used to undermine the evaluators and, therefore, their conclusions. The best way to stop this before it happens is to hire evaluators who behave in a manner that instills confidence in, and respect for their professional competencies. If there is a concern that there may be resistance to doing the evaluation among the project staff, this should be communicated up front to the evaluators and their strategies for dealing with such resistance should be discussed. For instance, in certain cultures it would be important for the evaluators to articulate their credentials to establish credibility with the project team. The evaluation manager can play a key role in this through regular check-ins with the evaluators and the project team.

"Whatever you say, ma'am"

It is important for an evaluation manager to be conscious of the politics of evaluation throughout the process. Recognize that evaluators are often consultants whose living is dependent on client satisfaction and referrals. As such, there are instances where pleasing the client may be put ahead of good practice. This is most commonly seen in evaluation reports that omit or play down problems that were identified, but it may also occur in the selection of methods – for instance, using a questionnaire because the client thinks it is the only credible data collection tool when qualitative information is actually what is sought.

To avoid this, evaluation managers should be sure to ask about professionalism during the reference check. It also helps to communicate openly with the evaluators that the organization seeks to learn not only about what they do well, but also about the areas in which they need to improve. An evaluation report that only addresses the positive but not the negative (or vice versa) will not be deemed satisfactory. At the same time, the evaluation manager should communicate the same message to the project team.

An evaluation report that only addresses the positive but not the negative (or vice versa) will not be deemed satisfactory.

"Being all things to all people often means being nothing to anyone"

There are many different stakeholders in peacebuilding projects – participants, project staff, the organization, partners, and donors – all of whom have different information needs when it comes to an evaluation. A common pitfall is to identify all these groups as the evaluation audience. While a donor may want to know if the organization being evaluated met the donor agency's mandate, the project participants may be most interested in sustainability or unintended negative effects. The project team on the other hand may want feedback on the process utilized. When making the key evaluation decisions at the project design stage, it is important to think through the audience question. Determining who will be the "users" versus the "readers" of the evaluation is effective in identifying the primary audience. See page 108 in the Evaluation Preparation chapter for more information.

"We'd be happy to do that for you"

Project staff members commonly draw up interviewee lists and offer to set up the interviews for an evaluation. This can appear to the evaluators to be a great way to save them from the tedious job of arranging appointments. However, project teams may select people who they think will speak most positively about their work or, when arranging the interview, they may intentionally or unintentionally tell the person what to say to the evaluator.

Consider, for instance, the Sudan IDP-host community example used in the Evaluation Preparation chapter on page 109. Bias or favoritism might become a factor if the project team handpicked those individuals whom they felt had gotten the most benefit from the project and in arranging an interview said, "Be sure to tell the evaluators how important this was to you and how much this changed your life. We really hope to get more funding and we only will receive it if the evaluation is good." Before the evaluation even arrives, the individuals have been told what to say.

In some cases, having project staff do these tasks may be necessary if the budget and/or time is limited. In these situations, the evaluator should identify the generic types of people to interview and then provide the project staff with a script or guidelines for their discussions with the potential interviewees. In addition, the evaluator should verify that there is a wide spectrum of people on the interview list and allot extra time for spontaneous conversations when she/he is actually on the ground.

"We want Wonderwoman/Superman"

For complicated or multi-faceted evaluations, it may be difficult to find one person who possesses all the skills and experience needed. One strategy for addressing this difficulty is to devise a team on which all the required skills are distributed among the different team members. To ensure the effectiveness of this strategy, it is critical that the team structure itself and the evaluation process to capitalize on the diversity of expertise among the members. There are four options within this strategy:

Conflict situations are highly volatile and dynamic; therefore, programming often needs to shift or, at times, change focus entirely in order to remain relevant to the needs on the ground.

- Having daily debriefs to do real-time analysis of information
- Jointly constructing the questions in the research instruments to ensure they include all key perspectives
- Identifying key issues from the various perspectives represented on the team during the analysis phase
- Joint report writing

"That logframe is so out of date"

Conflict situations are highly volatile and dynamic; therefore, programming often needs to shift or, at times, change focus entirely in order to remain relevant to the needs on the ground. Very often the program content is modified while the original logical framework is left untouched. Changing the content without changing the logical framework can create a dilemma for the evaluators because many evaluation design decisions

are based on the logical framework. It can create even greater problems if the evaluation objectives include results identification or adaptability of change since the results stated in the logical framework will no longer be relevant to the new work on the ground.

The easiest way to avoid an outdated design tool is to use them as they are intended to be used. As such, when the conflict situation shifts, use the design tool to assist in the creation of the new focus. If it is too late and the evaluation is about to start, the development of a logical framework that is representative of the new programming can be included in the terms of reference. This additional task would then need to be factored into the budget, timing and skill set of the evaluation team, yet it can contribute another useful deliverable to the program in addition to the evaluation. It is recommended that the program team be involved to some degree in the creation of the logframe to ensure its accuracy and utilization by the team beyond the evaluation.

Because conflict transformation projects often deal more in the realm of qualitative changes, such as decreased feelings of hatred or changes in attitude toward the "other," a quantitative-qualitative blend is best suited to grasp the difference the program is making in its entirety.

"It's not a result unless it's a number"

Traditionally, results have been considered credible if they were generated through quantitative research methods. This approach can only tell part of the story and it leaves many questions unanswered. For example, a questionnaire consisting of multiple choice questions may provide the project with information on how many people attended a town hall meeting and how much they learned from it, but it would not be useful for conveying any of the unique circumstances that brought them there or that caused them fear or concern in attending. This quantitative focus is partially due to the cross-over from the development field where evaluations of infrastructure or public health projects were handled almost exclusively in quantitative form.

Because conflict transformation projects often deal more in the realm of qualitative changes, such as decreased feelings of hatred or changes in attitude toward the "other," a quantitative-qualitative blend is best suited to grasp the difference the program is making in its entirety. This expectation should be communicated in the terms of reference and should be clearly visible in the evaluation plan.

"I'll just be a fly on the wall"

Some project staff may be very interested in the evaluation process and/or in what people have to say about the intervention and, consequently, may request or even insist on shadowing the evaluation team. Pending the evaluation objectives, identified use and team makeup, this interest may not be a problem and, if planned for, could be a capacity-strengthening exercise. However, in other cases it can cause difficulties.

Staff members who sit in on interviews or take part in the evaluator's debriefs can hinder the openness of these discussions. In some instances, staff presence can change the entire dynamic of an interview, to the point that the interviewee alters her/his story to ensure ongoing NGO support. Moreover, if the audience or use of the evaluation requires that there be a high degree of credibility in the methodology, permitting staff attendance in certain aspects of the evaluation may reduce overall confidence in the evaluation.

The easiest way around this pitfall is to include in the terms of reference development a discussion on the role of staff. That role must then be communicated clearly to all project staff. It should be noted that, if staff insist on accompanying the evaluators, a professional evaluator should immediately communicate the potential pitfalls to the evaluation manager.

"But I thought we were on the same page"

Since the fields of evaluation and conflict transformation are still figuring out exactly how they best fit together, a set of commonly held norms or best practices for evaluation have yet to be established. As a result, any assumption that the evaluator and the organization are on the same page regarding what constitutes "good practice" for the evaluation is a pitfall waiting to happen. One cannot assume, for example, that meeting with the project team to discuss the terms of reference or doing a debrief before leaving the country will take place unless it is specifically stated in writing.

Staff members who sit in on interviews or take part in the evaluator's debriefs can hinder the openness of these discussions.

To avoid disappointment, organizations should be explicit about their expectations of the evaluation process. It is best to document these expectations and include them as part of the contract. If that is not possible, then at a minimum, have a conversation with the evaluator prior to each stage of the evaluation process about what you as the client consider to be the norm. Some of the areas to consider are the alignment with organizational language, key steps in the process, ethics standards, report content, and expectations and confidentiality.

To avoid disappointment, organizations should be explicit about their expectations of the evaluation process.

Further Reading:

ALNAP Training Modules for the Evaluation of Humanitarian Action
Module 3: Managing and Facilitating Evaluations of Humanitarian Action, July 2003. http://www.alnap.org/training.html

The Evaluation Centre, Evaluation Management Checklists. http://www.wmich.edu/evalctr/checklists/checklistmenu.htm#mgt