Further Reading

Johanna Birckmayer and Carol Weiss, *Theory-Based Evaluation in Practice: What Do We Learn?,* Evaluation Review, SAGE Publications, Vol. 24, No 4, August 2000.

D. Fetterman, S. Kaftarian and A. Wandersman (eds.), *Empowerment Evaluation, Knowledge and Tools for Self-Assessment and Accountability*, SAGE Publications, 1996.

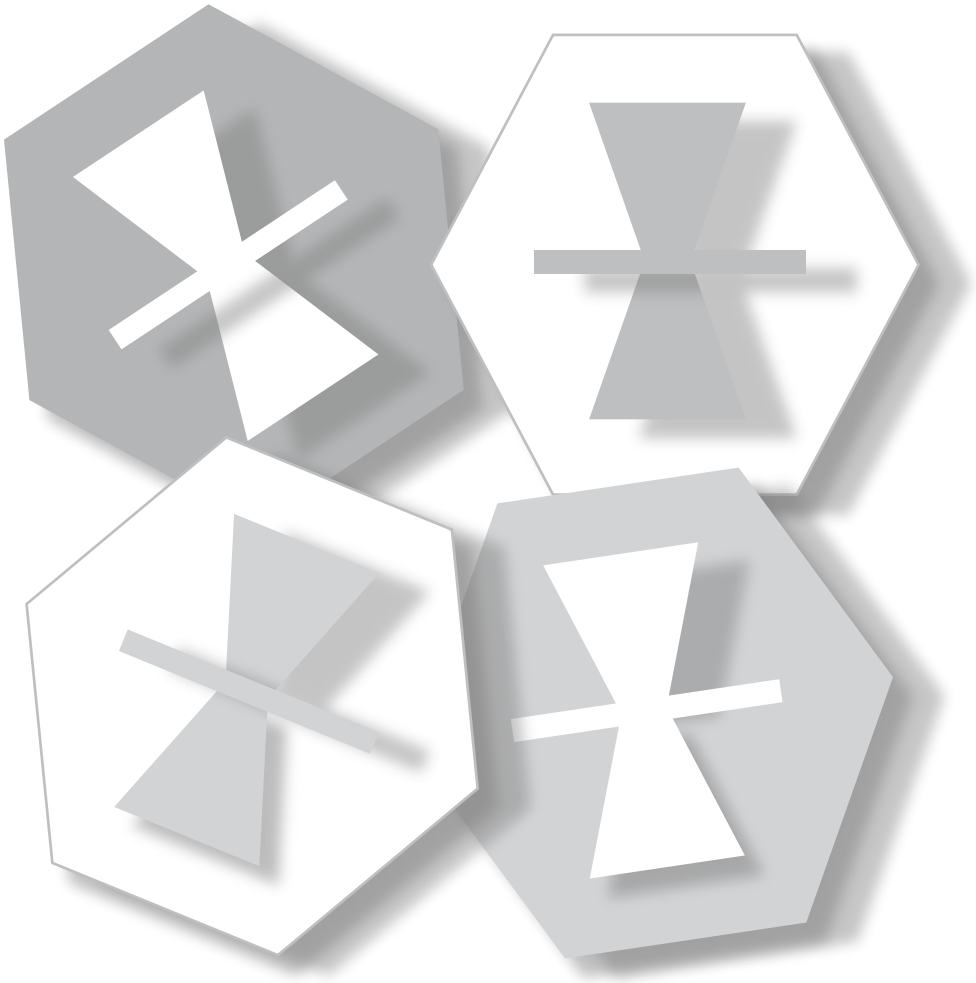
Michael Quinn Patton, *Utilization-Focused Evaluation: The New Century Text, 3rd Edition,* SAGE Publications, 1996.

Jay Rothman, "*Action Evaluation"* in *Beyond Intractability*, Guy Burgess and Heidi Burgess (eds.), October 2003. http://www.beyondintractability.org/m/action_evaluation.jsp.

Michael Scriven, *Pros and Cons about Goal-Free Evaluation* in Evaluation in Education, W.J. Popham (ed.), McCutchan Publishing, 1974, pp. 34-67.

Swiss Agency for Development and Cooperation, *Manual on Self-Evaluation,* February 1996. http://www.deza.admin.ch/ressources/deza_product_e_118.pdf.

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?

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. 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.

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

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

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

- 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.

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
    - o Appendix A – Terms of reference
    - o Appendix B – List of those interviewed (if not confidential)
    - o 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.

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.

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.

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.

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 or der 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

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.

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.

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.

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

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.

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.

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.

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.

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.

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

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:

- 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.

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.

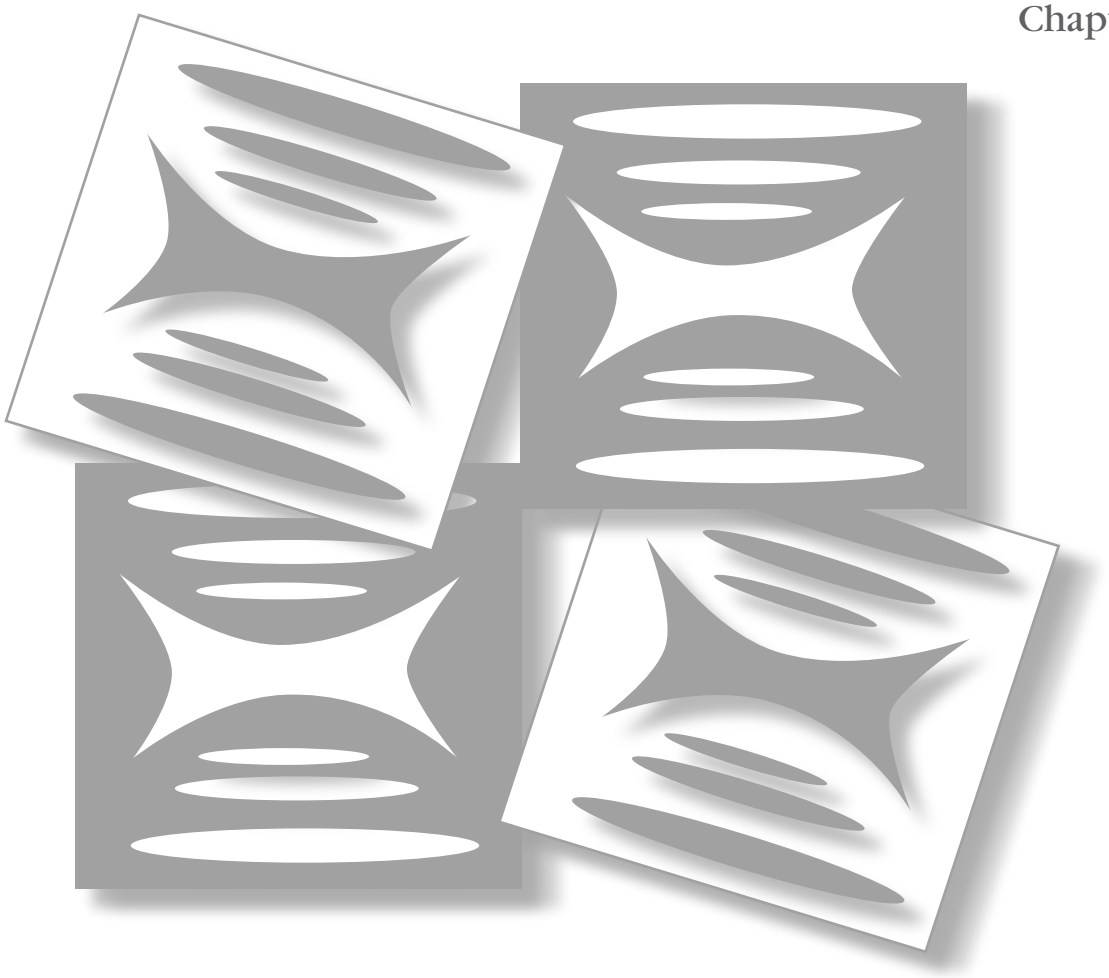
To avoid disappointment, organizations should be explicit about their expectations of the evaluation process.

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

Chapter **10**



STAGE THREE

EVALUATION UTILIZATION

This chapter contains:

1. Three steps to using evaluation conclusions and recommendations
    - Reflect and Generalize
    - Apply
    - Share: Internally and Externally
2. Evaluation Utilization Checklist

INTRODUCTION

"Creative thinking may mean simply the realization that there is no particular virtue in doing things the way they always have been done."

- RUDOLF FLESCH

In fact, the final report represents the completion of only the first two-thirds of a process whose main benefits are packed into the final third.

One of the most common misconceptions about evaluation is that the process is finished when the final report arrives on the project leader's desk. In fact, the final report represents the completion of only the first two-thirds of a process whose main benefits are packed into the final third. Imagine that the evaluators have brought a full bucket of water to the project leader on a scorching hot day. Someone has to reflect on how the water will be used and then take action to pour it into glasses, drink it, use it to water the plants, or bathe with it, otherwise it will just sit there and evaporate until it is no longer useful to anyone.

The benefits of utilizing evaluation results are multiple and they are at the heart of the effort to advance the peacebuilding field.

The latter third of the evaluation process focuses upon the use of the recommendations and conclusions. This stage is often referred to as the feedback loop, and it generally has two facets: internal learning and external sharing.

With the pressure to write proposals, implement projects and report to donors, internal learning may feel like an additional burden to the already overburdened practitioner. Yet the benefits of utilizing evaluation results are multiple and they are at the heart of the effort to advance the peacebuilding field. Within the project team, evaluation results enrich the discussion on whether the work is helping those it is supposed to help in the way it was planned. More pragmatically, it can indicate ways and means of improving the process. In addition, the advancement of models and designs indicates a level of expertise that can be compelling to donors.

As conflict transformation theory and practice continue to develop, a commitment by organizations to internal learning and external sharing has great potential to effect change within the practice of this field.

Sharing the evaluation results and the subsequent reflections and adaptations to the project externally increases transparency between the organization and its stakeholders and donors. Such sharing can also foster dialogue and reciprocity between organizations. It contributes to the growth and value of the field while reducing the amount of resources wasted through reinventing the wheel. As conflict transformation theory and practice continue to develop, a commitment by organizations to internal learning and external sharing has great potential to effect change within the practice of this field.

When do I start thinking about the "use" of an evaluation?

Begin thinking about how the evaluation will be used during the preparation stage. The evaluation objectives and how the project team hopes to use the information generated by the evaluation must be aligned. If, for instance, the evaluation objective is management and administration but the team hopes to use the evaluation information to understand more about the effects of the project, the team is unlikely to find the findings useful. External use of the evaluation should also be addressed briefly during the preparation stage when deciding on the evaluation's audience and its potential readers.

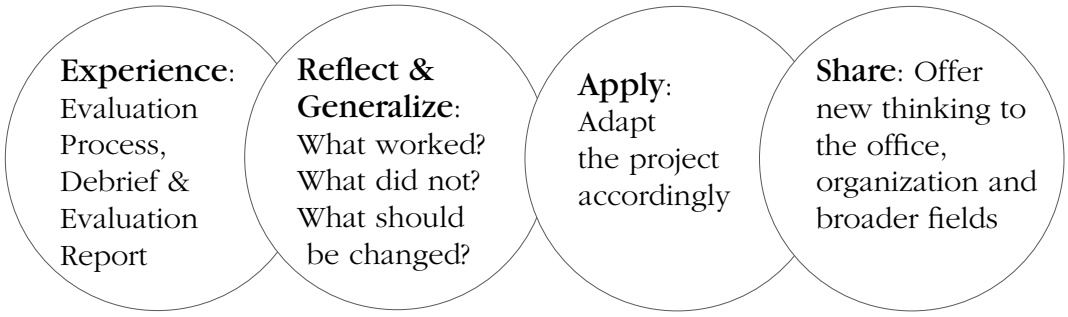
Of course, the evaluation implementation itself is a learning experience, but the real attention to internal learning starts when the evaluation team presents their draft conclusions while still in-country. A smart evaluation manager will remind the project team at this stage that there are more steps left to come.

The evaluation objectives and how the project team hopes to use the information generated by the evaluation must be aligned.

What are the steps needed to use the content of the evaluation report?

There are a series of overlapping steps, based on the Kolb adult learning stages outlined in the Learning Chapter page 6, that should be taken to maximize the internal and external benefits of an evaluation report. First, the project team goes through the experience of the evaluation – preparation, evaluation plan, process, debrief and review of the final report. The team then reflects on the conclusions and recommendations, and generalizes the information beyond the immediate project. From there the focus shifts to how this newly acquired knowledge should be applied to the project or beyond. Finally, the team considers what needs to be shared with other teams, with the broader organization, and with the field as a whole.

Applying Kolb's Theory of Learning



As illustrated in the diagram, these steps overlap. What is important is that all the steps be taken and that they be completed in real-time. In other words, the process should not be delayed until the team thinks it will have more time to discuss the results.

Well-intentioned evaluation managers and project teams sometimes jump immediately to application. This is not altogether bad; however, these discussions often become stuck in the technicalities of the project and may not offer the opportunity to reflect on the macro-level issues or to address issues at a generic level separate from the immediate demands of the project.

EXPERIENCE: The evaluation has been planned, data collected, and tentative conclusions drawn. Both the in-country debriefing and the discussion with the project team have taken place. The remaining steps begin at this point.

REFLECT AND GENERALIZE: This step involves consideration of the findings from the perspective of the immediate project and long-term learning, review of the recommendations, and identification of the actions to be taken.

As with a conflict resolution process, simply gathering the right people is not enough. The process requires additional thought and creativity.

Once the evaluation is submitted, someone – often the evaluation manager – should be appointed as the learning facilitator. Remember, this role may not be necessary if it is already part of the evaluator's responsibilities. The learning facilitator works with the project team to identify who should be included in these reflections. Together they should consider people located horizontally from the team as well as vertically above and below the team decision makers. Horizontally, staff members responsible for similar processes, partners, and stakeholders should be considered since they can benefit from the experience of others. Vertically, supervisor(s) – whether immediately above or even more senior – should also be included.

Next, develop a facilitated process to support the reflection and generalization. As with a conflict resolution process, simply gathering the right people is not enough. The process requires additional thought and creativity. For larger groups, this may involve a multi-stage process or different processes for different groupings. Sometimes it is helpful to focus solely on the evaluation conclusions in the first conversation and to leave the recommendations for the next discussion. This allows the project team to focus on the findings and to have a rare opportunity to discuss assumptions, principles and techniques outside the context of the day-to-day pressure of implementation.

In designing the process and selecting the participants, the conflict setting should not be ignored, particularly if staff members are from the area of conflict.

In designing the process and selecting the participants, the conflict setting should not be ignored, particularly if staff members are from the area of conflict. Ensure that the process does not permit individual staff members, groups in the community or aspects of the governance structure to be blamed either for problems that arose or for a lack of results. Affirm the challenges of working on conflict issues and recognize the need to work around them.

Whatever the process used, the discussion should facilitate reflection and generalization. Consider questions that range from the conceptual to the technical. The discussion need not be restricted to the evaluation results; however, be careful that the conversation does not avoid the thorny issues that an evaluation may raise. Depending on the evaluation objectives, the report conclusions may bring up some or all of the following questions for the project team to discuss:

- Did our theory(ies) of change work? Why or why not?
- Would a different theory of change have catalyzed a greater transformation?
- Was the context analysis accurate?
- Did our strategy link to the analysis in the best way?
- What happened as a result of these activities?
- What did not happen – that we expected to happen – as a result of these activities?
- Was the process we used within each activity the right one?
- How can we improve our techniques (e.g., training techniques)?
- What do the evaluation recommendations mean for the project and organization?
- What are the key learning points from this for the future?
- What were the critical factors for success?
- What are the pitfalls to look for next time?

To facilitate sharing and building institutional memory, the process and conclusions from these discussions should be documented.

Key Principles of Learning [21]

- Help others actively interpret information so that they can construct new knowledge for themselves, rather then relying solely on paper formats.
- Situate abstract tasks in real contexts so that the relevance of the task is apparent and others can adopt the new knowledge.
- Separate knowledge from a single specific context in order to maximize knowledge sharing.
- Provide others with many examples of a new concept.
- Explain how essential features of a new concept are reflected in a variety of different settings.

[21] Adapted from the *UNDP Handbook on Monitoring and Evaluation for Results*

Identifying what did or did not work is an important first step. However, without an intentional discussion about what will be changed both in the immediate aftermath of the evaluation and in the longer term, it is unlikely that reflection alone will be enough to change the way projects are designed and implemented.

Begin by determining if everyone involved in the reflections conversation needs to be involved in setting action steps. This is facilitated by keeping the discussions on reflections and generalizations separate. Then review the team conclusions that resulted from the reflections and generalization discussion and determine the action steps to be taken to implement those conclusions. This discussion should also address which of the evaluation recommendations will be adopted. It is important to note that a project team is not required to act on every recommendation. The team should reflect, however, on all the recommendations and determine which will be accepted or rejected and why.

For those recommendations not adopted, the reasons for not adopting them should be stated.

APPLY: This step moves the team back into action. To facilitate this, it may be helpful to develop a utilization plan. Action steps based on evaluator recommendations should be noted in conjunction with the recommendation from which they are taken. For those recommendations not adopted, the reasons for not adopting them should be stated. The utilization plan should also outline what will be done, by whom, and when.

This document can be a valuable addition to institutional memory since it captures decisions that may be useful for future evaluators, proposal writers, donors or new staff joining the team. Pending the content of the utilization plan, it can also be beneficial to attach it to the evaluation circulated to donors and key stakeholders. This shows that the organization is taking the evaluation seriously and has the capacity to learn and improve.

It may be helpful to extend the learning facilitator's role by a few months or even a year so that this person can check in with the team to ensure that the utilization plan has been implemented. As time passes, the learning facilitator can also initiate discussions on what difference those changes have made to continue the learning process.

SHARE: INTERNALLY Finally, what was learned needs to be shared with others both within the office and throughout the organization. At a minimum, circulate the evaluation and utilization plan to relevant parties such as department managers, senior personnel, field directors, etc. Other actions may include:

- Circulating the evaluation and utilization plan to wider range of staff

- Hosting a workshop on the results and lessons learned
- Adding panels to regional staff meetings on lessons from the evaluation
- Developing a short lessons learned summary
- Contributing a paragraph on the two critical factors for success in the internal newsletter
- Convening meetings to share the learning with regional or technical staff who support a wide range of programs
- Posting the evaluation on the organization intranet

Consideration should be given to the process used to share information if the goal is not only to assist people in knowing it but also applying the new information. Use the Key Principles of Learning depicted on page 182 to inform design of these mechanisms.

"I can honestly say that not a day goes by when we don't use those evaluations in one way or another."

Written by *M. M. Rogers* and illustrated by *Lawson Sworh*

SHARE: EXTERNALLY In addition to the learning benefits from the evaluation results, a wise organization reaps an array of valuable benefits from distributing the evaluation externally. Different formats, such as key results document, may need to be developed for different external audiences in order to be effective. Possible external uses include:

Donor Relations

- Incorporate evidence-based results to strengthen proposals
- Use the document as a means of outreach to new donors
- Strengthen established relations through discussions about the results and the changes the organization is making as a consequence of the evaluation

Public Relations

- Host a meeting with interested stakeholders to discuss the results
- Include key quotes on websites and brochures
- Add a "Results" or "Accomplishments" section to the annual report
- Include the results in a key talking points packet for the press
- Produce an "Accomplishments" brochure that focuses on the difference the work has made in the world
- Write a concise summary and circulate to electronically to relevant peacebuilding listservs

Academia

- Write journal articles that include the evaluation results or that are based on the evaluation experience
- Invite academics and researchers to use the data collected for the evaluation in their studies
- Present papers on the evaluation at conferences

Within the fast-paced context of a conflict situation, engaging in systematic reflection can prove to be extremely challenging as the needs or opportunities that prompted the project in the first place may no longer be relevant.

I'm really busy. Can all of this be put together in a checklist?

Within the fast-paced context of a conflict situation, engaging in systematic reflection can prove to be extremely challenging as the needs or opportunities that prompted the project in the first place may no longer be relevant. In addition, a sense of urgency can develop that one needs to learn new things to address up-and-coming problems. If this sense dominates, the ability of the peacebuilding field to improve

and to increase its effectiveness in achieving its goal of building positive peace around the world will remain limited. Moreover, project teams and organizations will not learn to do their work better, which will also limit the sharing of knowledge within the field.

Use the following Evaluation Utilization Checklist as a tool to ensure that none of the steps are missed. The "Lead Actor" column indicates who is responsible for the task, while the "Who is involved?" column covers key participants within the organization. The "When" column requires a sense of timing for the task.

Evaluation Utilization Checklist

Tasks	Lead Actor	Who is involved?	When
Determine evaluation objective(s)			
Conduct draft conclusions and debrief with project team			
Determine who will be involved in reflections conversation			
Develop process for reflections conversation			
Circulate final evaluation report to all involved in conversations			
Conduct first reflections conversation			
Document thoughts and ideas from conversation			
Determine who should be involved in making decisions about changes at the project, program and/or organizational level			
Identify adaptations to be made including responses to the evaluation recommendations			
Development of utilization plan			
Evaluation and utilization plan (including new knowledge) circulated to relevant staff			
Monitor how the new learning and utilization plan has been applied			
Incorporate results into organizational working knowledge through a variety of forums such as • Workshops • Internal Newsletter • Panels • Lessons Learned Briefing			
Consider use of results in donor relations, public relations and academia			
Circulate to other organizations in the field			

Advanced Concept

Utilization from a long-term perspective

Over time, an organization will accumulate more and more evaluations, which will create opportunities for greater utilization, improvement in evaluation quality, and the maximization of learning. Some options to consider include:

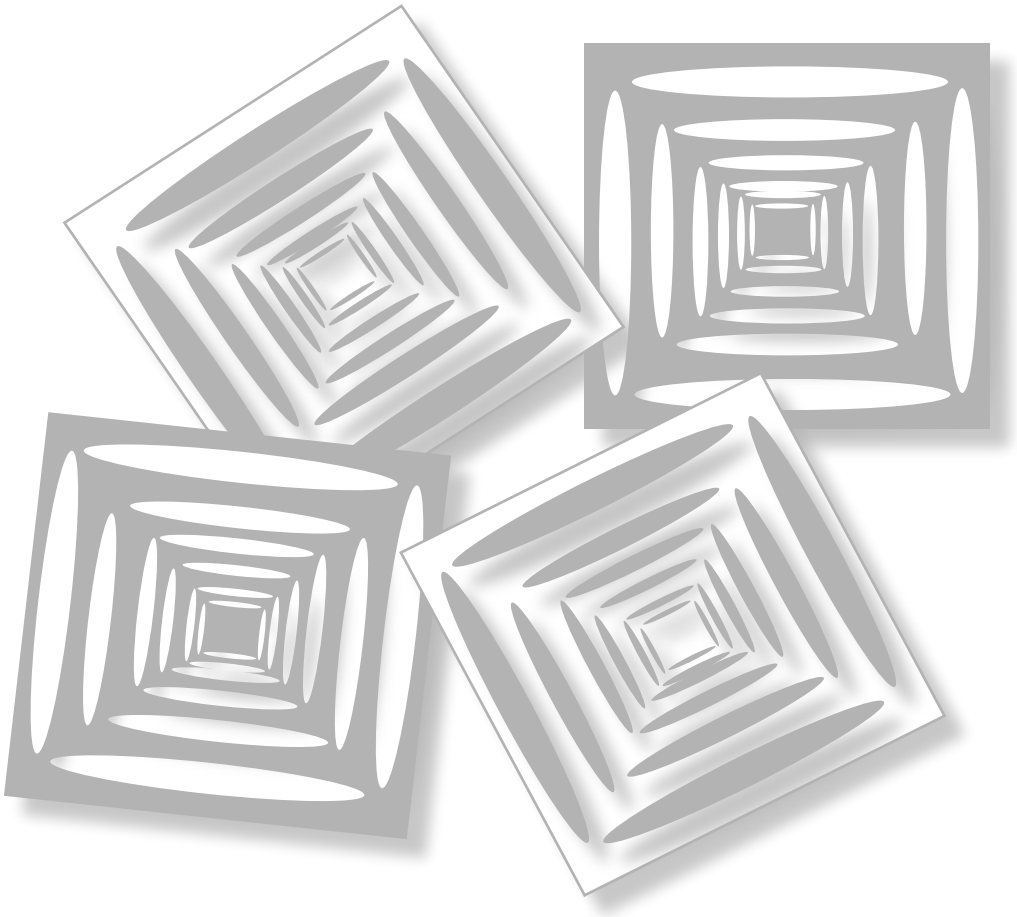
- ***Evaluation Synthesis:*** *The gathering and summarizing of the results of evaluation studies of similar programs (by theme or tool).*
- ***Standards of Practice Guidelines****: Development of standards of practice or guidelines that evaluations can use to assess programs against.*
- ***Meta-evaluation****: Assessing existing evaluations against pertinent standards to help improve evaluation implementation in the future. This is a form of evaluation quality control.*

Further Reading:

Michael Quinn Patton, *Utilization-Focused Evaluation*, Third Edition, SAGE Publications, 1996.

David A Kolb, *Kolb's Theory of Adult Learning*, 1984.
http://www.learningandteaching.info/learning/experience.htm

Chapter **11**



In Design, Monitoring And Evaluation For Conflict Transformation **ETHICS**

This section includes:

1. Discussion of unethical practices

2. Categories and strategies for dealing with common ethical issues
   - *Protection of people*
   - *Freedom from political interference*
   - *Quality data collection techniques*

3. What is informed consent?

4. Internal versus external evaluators and ethics

INTRODUCTION

"To enjoy the things we ought and to hate the things we ought has the greatest bearing on excellence of character."

- ARISTOTLE

This chapter discusses ethical issues in design, monitoring, and evaluation of peacebuilding projects. It begins by offering guidance on what constitutes an unethical practice and how such practices might arise at each stage of the project cycle. Not surprisingly, the ethical issues related to baselines and evaluations overlap in a number of areas, whereas the design stage has several unique ethical challenges of its own. Within each section, practical strategies for preventing and avoiding unethical choices are offered. The practice of informed consent is considered in relation to the realities of evaluation of peacebuilding. Finally, the chapter explores the different ethical issues faced by internal and external evaluators.

What is an unethical practice?

Not knowing what constitutes best practice is incompetence. Knowing what best practice is, but not knowing how to achieve it, may be inexperience. Knowingly not following best practices, when one knows how to achieve it, is unethical.[22]

There are many types of unethical practice in design, monitoring, and evaluation of conflict transformation programs. On one hand, there are ethical issues that commonly occur in DM&E of peacebuilding that generally have a "right" and a "wrong" answer. Changing data to represent a project in a more positive or negative light, for instance, is clearly wrong.

Conversely, an ethical dilemma often does not have a clear right or wrong answer, and because of this ambiguity, decisions should be considered on a case-by-case basis. The cultural norms, values, and experiences of those involved often play a significant role in the decision. An example of such an ethical dilemma can be seen in the decision about whether or not to insist on the equality of youth voices in decisionmaking as part of a participatory evaluation in a society that honors elders and the roles they play as key decisionmakers.

For some, what is discussed in this chapter may be viewed either as standard political actions to benefit one side or another or simply the lack

[22] Nick Smith, *An Analysis of Ethical Challenges in Evaluation,* American Journal of Evaluation, Volume 23, No. 2(2002)

of application of best practices. Regardless of the label applied, if the action does not fall within appropriate principles of conduct for design, monitoring, and evaluation, it fits within this discussion on ethics.

What are the ethical challenges common in the design stage?

When discussing the development of a peacebuilding project, innumerable ethical challenges may arise. The four challenges discussed below, however, are those that are likely to apply in most situations.

First, practitioners have an ethical obligation to involve the parties to a conflict in determining the changes that will satisfy their respective interests. Although people in conflict may not always be able see how to transform their disputes at the outset of discussions, they ultimately make the final choices once additional options have been explored.

Second, practitioners have an ethical responsibility to consider any possible negative ramifications that may occur as a result of a project and to do everything feasible to eliminate them. Consider, for instance, a youth project in Israel that targeted teenage boys, ages 13-15, from politically hard-line families. The goal of the project was to change attitudes from supporting violence to a recognition that there are multiple ways beyond violence to resolve the political situation. As the project progressed, some of the participants started to challenge the adult members of their families in political discussions. In one case, this led a father to physically assault his son as punishment for what the father saw as the son's disloyal and disrespectful opinions. Such an unintended negative effect might have been prevented if, in the design stage of the project, this scenario had been identified and preventive measures adopted. These measures could have included engagement with parents or the incorporation of techniques for dealing constructively with families about sensitive issues so that the participants would be prepared for such a situation.

Third, practitioners have an ethical obligation to develop projects that maximize the opportunities for change. This maximization is determined on a situation-by-situation basis, but it broadly encompasses creating change among the most people, in the fastest way possible, for the greatest possible positive change, and with the least possible negative consequences. The ethical challenge arises when project designs that do not maximize the opportunities for change are seen as easier to implement or have more readily available sources of funding.

Finally, the development of indicators can be an ethical issue. For example, there may be indicators that reflect changes of less importance or that signal changes on issues that are not directly affected by the project but which present the work in a more positive light than would an accurate indicator.

In an attempt to address these ethical challenges, organizations can add a set of questions to their project design processes that they routinely review. These questions might include some of the following:

- If there were absolutely no restrictions in terms of capacity, time, or funds, how would we modify this project?
- Have discussions with the prospective donor taken place to explore options that may be more productive or beneficial to the stakeholders?
- Have the connections between the analysis and the proposed project been explicitly outlined?
- Can the stakeholder's perspective be seen in the final design?
- What are the potential negative results that could occur as a result of this project? What steps need to be taken to minimize the potential negative consequences to participants, staff, or the community?
- Are there other programs currently operating to which this project should be connected in order to maximize results?
- If the team members could only implement one project, which would they select as the most important, and would it make a difference?
- Were other options fully discussed based on the conflict assessment, particularly those not part of our regular activities?

What are some of the common ethical challenges for baselines and evaluations?

There are a number of ethical issues and dilemmas to consider when implementing baselines and evaluations. In some cases, the same issues apply to monitoring as well. The ethical challenges can be grouped into three broad categories: protection of people, freedom from political interference, and quality data collection techniques.

1. Protection of People

The ethical challenges related to the protection of people can be subdivided into six major themes: avoiding personal duress, guaranteeing confidentiality, considering safety, setting realistic expectations, protecting the organization's credibility, and avoiding research subject fatigue.

This category (protection of people) spans all units of analysis, from the individual, to the implementing organization, to the target group as a whole (e.g., all Hutus or all ex-combatants).

AVOIDING PERSONAL DURESS: Data collectors should consider the potential negative consequences that could arise from delving into an individual's personal experience. For instance, silence can be a coping strategy for some victims of violent conflict; however, they may be asked to talk about their experiences as part of the evaluation process. This process therefore risks undermining the participants' coping strategy without offering the necessary support structure to provide assistance if it is needed.[23]

Evaluators should approach some groups in places of war, such as victims of rape or torture, with caution, and ideally consult with experts on the appropriate ways to engage with these groups, if at all. However, conflict zones are rife with individuals who have unhealed psychological wounds and trauma that are not apparent. Evaluators should therefore look for signals of duress, such as agitation or tears, in their subjects and be prepared to handle the situation appropriately. Seeking advice from experts on this issue, prior to data collection, is a prudent step for the professional evaluator.

GUARANTEEING CONFIDENTIALITY: It is important for individuals providing data for a baseline or evaluation, whether through surveys or in one-on-one interviews, to understand how their names will be used in connection with the information they provide. The evaluator must explain clearly how the information will be attributed in the final deliverable. In other words, will the person's name be used, along with her/his ideas, in a quote format or will attributes be used to provide a context for the comments (e.g., women in the village), or will the information simply stand alone?

In conflict settings, where speaking out against one's group or the government, for instance, may prove deadly, the norm is to guarantee confidentiality to all individuals who participate. In this case, not only does the evaluator need to explain that the data is confidential to each individual, she/he must do preparatory work to ensure that confidentiality can be guaranteed. More on this issue can be found later in the chapter under Informed Consent, page 198.

Special care is due when writing the baseline or evaluation report once confidentiality has been promised. In local settings or where people are assumed to hold particular views, even general attributes in connection with specific statements may be identifiable by the community. For instance, if there are only ten positions on the district council and six individuals have held their positions for years, attributing a statement to a new member to the district council is almost the same as using the

[23] Cheyanne Church and Julie Shouldice, *Part II.*

person's name. This can become a difficult dilemma when the identity, position, or standing of a person inherently provides insights on her/his statements or opinions.

CONSIDERING SAFETY: In conflicts where communities are segregated, like Kosovo or Northern Ireland, being seen speaking to an outsider can be enough to cause suspicion within an individual's community. A number of questions might result: Who was the stranger? Why were they here? What did they want to know? What did you tell them? In a calm situation, suspicion may end with harmless gossip; however, if tensions rise or are already high, the suspicions could grow into more serious outcomes for the individual, such as expulsion or physical harm. Evaluators have an ethical responsibility both to consider the safety of the individuals who provide them with information and to plan their data collection efforts to minimize any possible risk. Where meetings take place, who introduces the evaluator to the individual, and who should be told about the evaluation and the purpose of the visits are all important considerations for an ethical evaluator.

Another safety dilemma can arise when an evaluation team hires members from the conflict setting. The members from the community may have greater freedom of movement in areas experiencing active conflict and often travel alone to these locations to collect data. What is the team's responsibility for the personal safety of these local team members when they enter high-risk areas? The dilemma lies in what constitutes too much risk. It may be useful to consider the following rule of thumb: If the team member from the area would not travel to the conflict area independently, regardless of the foreign vehicle or official trappings, other options should be considered for accessing the data being sought. (See the Evaluation Management chapter, page 168, for more information on accessing data in situations that are too dangerous to enter.) In addition, do not assume that a team member from the area is aware of the security concerns at the time the work is to be done.

SETTING REALISTIC EXPECTATIONS: Anyone collecting data for a baseline, monitoring, or evaluation should be careful not to set undue expectations during the data collection process. It is often the case that, in an effort to express thanks, generate excitement, or convince people to answer questions, the data collector inadvertently raises expectations unrealistically. Consider the following example.

A practitioner was monitoring the progress toward results of a program seeking to increase a community's knowledge of the city's grievance procedures. He wanted to speak to community members who were residents of government-subsidized housing in an area of violent crime because the project team felt that the data should be disaggregated for socio-economic standing. (See the Methods chapter, page 216, for more information on disaggregated data.)

Convincing residents of the housing project to speak with him was difficult, however. In an effort to entice more people to cooperate, he opened the conversation with the following statement: "If our group is going to be able to help people deal with crime in this area, we need people to answer a few questions." This led many of the residents to conclude that the NGO was going to work actively in their community to decrease crime. Arguably, knowing more about grievance procedures might help in this regard; however, the NGO's intention was never to directly address crime and violence. As a result, it set unrealistic expectations.

PROTECTING THE ORGANIZATION'S CREDIBILITY: Those collecting data should also be aware that their actions are often deemed to be an extension of the organization being evaluated. Consequently, if a team member behaves inappropriately it can harm the organization's reputation, and – in more serious cases – the inappropriate behavior may derail any progress achieved to date from the work.

Consider the following example from Northern Ireland. "[A]n evaluator… enters a tense conflict situation to evaluate a cross-community dialogue project with leaders of opposing communities. The evaluator is permitted to meet with the participants in the program because of the goodwill and trust established between the conflict parties and the implementing agency. However, if the evaluator does not operate within the norms of the communication established by the agency such as meeting with an equal number of representatives from each side, or is interpreted as being biased by one of the parties, this can severely damage the agency's credibility with the parties and constrain the dialogue process."[24]

AVOIDING RESEARCH SUBJECT FATIGUE: In areas where a great deal of research is done, there can be problems with beleaguered research subjects being asked repeatedly to offer information on similar themes in relatively short periods of time. Not only does such repetition steal valuable time from the individual, it also dilutes the authenticity of the answer since the person has been asked about the same subject so many times that her/his response becomes almost "pre-recorded."

Organizations considering an evaluation would therefore be prudent to ask other agencies working in the area or their donor if other evaluations are pending. Sometimes it is possible to combine evaluation or baseline projects. This not only shows respect for the research subjects but it can also decrease the cost of the research to the organization. Furthermore, professional evaluators should inquire about the possibility of cooperation with other organizations at the earliest stage possible within the evaluation since there may still be time to combine research with other efforts.

[24] Church and Shouldice, *Part II.*

2. Freedom from Political Interference

Evaluations should be conducted free from political interference from the implementing organization, the donor, the evaluation team, and the stakeholders. There are many ways, and many different reasons why, political pressure might be applied to an evaluation process. Implementing organizations may see the evaluation as a way to promote themselves to donors and, as a result, will steer the evaluator toward only those people who will speak positively about the organization. Donors may view an evaluation as a way to justify a decision to end funding to a sector or organization by requiring a methodology that misses many of the positive results. The evaluation team may wish to secure ongoing contracts with the implementing agency by presenting the agency in an unearned positive light. Stakeholders may see the evaluation as the only way to access additional resources for their community and, therefore, they may lobby the evaluator to make specific recommendations.

Evaluations provide far more opportunity for political interference than do baselines, though baselines are not exempt from meddling. Some of the more common political interferences can be found in the table below. An "X" indicates whether the political interference applies most commonly to baselines, evaluations, or both.

Common Political Interference in Baseines and Evaluations

Common Political Interference	Baseline	Evaluation
Implementing agency pressures the evaluation team to omit weaknesses from the final evaluation report.		X
Staff members ask the evaluation team to show them in a positive light.		X
Donors require the evaluator to use a methodology that is not optimal for the information being sought.	X	X
Staff pressure participants into being part of the study.	X	X
The implementing agency or donor already has an answer and writes the terms of reference in a way that sets up the evaluator to justify that answer.		X
Staff members coach project participants on the kinds of responses they want given to the evaluator.		X
Evaluators are pushed toward specific sets of people who are unusually positive or negative.	X	X
Not providing the evaluation team with reports that capture concerns or negative effects of the project.		X
Not including on participant lists those who have dropped out of the project.		X
Creating documents such as reports or project logs to meet the evaluation team request during the evaluation.		X

There are a number of strategies that can be adopted when there is political interference. Some of the most common strategies recommended by professional evaluators include:

FRAMING THE ISSUE OF CONCERN: Consider the attempted interference as a regular part of negotiations rather than as an unmanageable impediment to a quality evaluation process.[25] In this case, the evaluator should reframe the concern as an issue that requires additional negotiation with the party exerting the political interference. This technique is often combined with Communication & Education, which is described next.

COMMUNICATION & EDUCATION: Ensuring that all stakeholders understand the steps in the process and the rationale behind them will decrease opportunities for misunderstanding and potential malpractice. In the first stage of the evaluation, explicitly develop the principle of transparent communication, whereby all parties provide explanations of their requests and choices, and provide opportunities for discussion.

DETAILED & DOCUMENTED PLANNING: An evaluation plan and terms of reference that are well-thought through and documented can be very helpful for at least two reasons. First, they offer clear boundaries and decisions that are less open to interpretation. Second, they provide an historical reference that the evaluator can refer to later if inappropriate pressure arises during the evaluation process. A detailed contract with clearly defined grievance procedures can also be helpful if disagreements arise.

TIMELINESS: When actions occur or statements are made that seem intended to exert undue political pressure, they need to be faced immediately and directly.

INCREASE THE SEATS AT THE TABLE: The more stakeholders represented at the table, the more likely political interference will either not arise or will be handled in an appropriate manner.

CHECKS & BALANCES: Having an evaluation team rather than an individual evaluator can provide checks and balances when ethical challenges arise. In addition, if political pressure is at the heart of an issue, there is always more strength in numbers (i.e., as opposed to an individual evaluator on her/his own). Another good check and balance to put in place is an independent evaluation manager. See the Evaluation Management chapter for more information on the role of the evaluation manager, page 137.

CONSULT EXPERTS: If there is any uncertainty as to what is acceptable, consult an expert. If the organization has internal DM&E expertise, check to see if there are any norms that the organization has chosen

[25] Nick Smith, *An Analysis of Ethical Challenges in Evaluation,* American Journal of Evaluation, Volume 23, No. 2 (2002).

to utilize or request guidance. If there are no internal resources, use evaluation network listservs to request input or contact academics in the evaluation field. If the question is brief, people are generally happy to provide their input.

ADHERENCE TO AND DISCUSSION ABOUT PROFESSIONAL PRINCIPLES: Professional associations increasingly issue principles or norms of ethical practice. National evaluation associations are no different; therefore, be sure to check if professional principles have been issued for the country in which the evaluation will be conducted. If none exist and/or there is no national evaluation association, a good alternative is the American Evaluation Association Guiding Principles below. These principles may need to be adapted somewhat to other settings, but they offer a useful point of reference for a discussion with stakeholders in an evaluation.

American Evaluation Association Guiding Principles

A. Systematic Inquiry: Evaluators conduct systematic, data-based inquiries about whatever is being evaluated.

B. Competence: Evaluators provide competent performance to stakeholders.

C. Integrity/Honesty: Evaluators ensure the honesty and integrity of the entire evaluation process.

D. Respect for People: Evaluators respect the security, dignity and self-worth of the respondents, program participants, clients, and other stakeholders with whom they interact.

E. Responsibilities for General and Public Welfare: Evaluators articulate and take into account the diversity of interests and values that may be related to the general and public welfare.

HONOR YOUR OWN INTEGRITY: If you feel you are being asked to do something that does not "intuitively" feel right, raise the issue with the organization before you proceed. Some may call this the "Can I sleep at night?" measure.

3. Quality Data Collection Techniques

In addition to using best practice in data collection to ensure sound and credible inputs for analysis, evaluators also need to consider some ethical challenges that can affect data quality.

One such challenge is whether or not to reimburse people for the time they have given to provide information. In many peacebuilding programs, the average participant would qualify as being a member of the

world's poor. Taking two hours of time to participate in a focus group rather than earn money or gather food may have a substantial effect on the person's livelihood. Should they be reimbursed for their time, with money, transport, or food? Classic social science research would state that they should not be reimbursed since remunerating people for their opinions may cause them to alter their statements or responses to be unduly positive or supportive of the topic.

However, the rules of social science methodology were created in the "developed West" and often need to be adapted to non-Western contexts. If possible, it is always better to avoid providing material incentives since they can potentially skew the results. Nonetheless, this is not always the most ethical stance and, in certain cases, creatively identifying ways to reimburse people for their time is appropriate. Offering lunch, funds for transport, or a small item such as a bucket are potential examples. If such items are offered, assure each person that the reimbursement is guaranteed regardless of the information offered.

The second dilemma to consider is the balance between respecting local customs and advancing an agenda the organization or evaluation team deems important to the project. One of the most widely known illustrations of this is gender inclusion. It is commonplace for evaluators to want to engage men and women in their data collection, yet in some situations, accessing women's opinions may be counter to local customs. This can be particularly true if the evaluators are solely male and wish to speak to women without the presence of local men. Should the beliefs of the evaluation team override local customs? In these situations, it may be best to turn to the implementing organization for guidance.

What is Informed Consent?

Informed consent is the process of educating participants in the research about the purpose, procedures, risks, benefits, and alternatives to participation. In social science research at the academic level, informed consent is a standard and required part of any research project. It is seen as an ethical obligation of the researcher and as a key part of the protection of the people involved in the study. In these cases, consent needs to be obtained in written form from participants before they become involved in the research. It is far more than simply obtaining a signature on the consent document. It is about the individual's understanding and willingness to participate in the study.[26]

This standard of written consent is not yet the norm in international conflict transformation evaluation. Complying with the written documentation requirement may never be feasible because of illiteracy as well as confidentiality and security concerns in conflict settings. The essence of informed consent holds true, however, regardless of

[26] Informed Consent Overview, Institutional Review Board, University of Minnesota, 1998 http://www.research.umn.edu.consent/modlsoc/modlsec1.html

the setting. Participants in an evaluation should be informed of the purpose, process, risks, and benefits of participation and be given the opportunity to decline to participate.

Are there different ethical dilemmas and issues for internal versus external evaluators?

Ethical challenges do not generally differ between internal and external evaluators. What does seem to differ is the way in which issues are resolved. One of the primary factors behind this difference is the different relationship to the organizational structure that each holds.

It is the nature of those relationships that is critical for considering ethical dilemmas. Internal evaluators are situated directly within the organization whereas externals are outside the organization and are related to many different entities at the same time. The organization that constitutes the most important relationship to an external evaluator is rarely the one she/he is evaluating. Yet for the internal evaluator, the focus of the evaluation – a project within her/his organization – is generally the most important relationship to the evaluator's professional position.

Ethical dilemmas therefore arise with people who the internal evaluator knows well and works for routinely. The internal evaluator often feels that she/he has fewer options in challenging situations. To foster a sense of belonging and long-term community, an internal evaluator may feel the need to be more conciliatory about challenging issues. External evaluators generally have more latitude because their connections to the group involve a particular project and they have been brought in for their expertise on that project.

This situation may mean internal evaluators are more vulnerable to poor practices exerted by the organization or donor which result from conflicting roles associated with being both a professional evaluator and a member of an organization. However, the personal relationships that consulting professionals develop with their clients, and the expectations engendered by clients' direct hiring and reimbursement of the professional, may also exacerbate ethical dilemmas. Due to the inherent power dynamic, it can appear against the consulting professional's best interest to pursue ethical norms that seem to conflict with a client's self-interest.

Further Reading

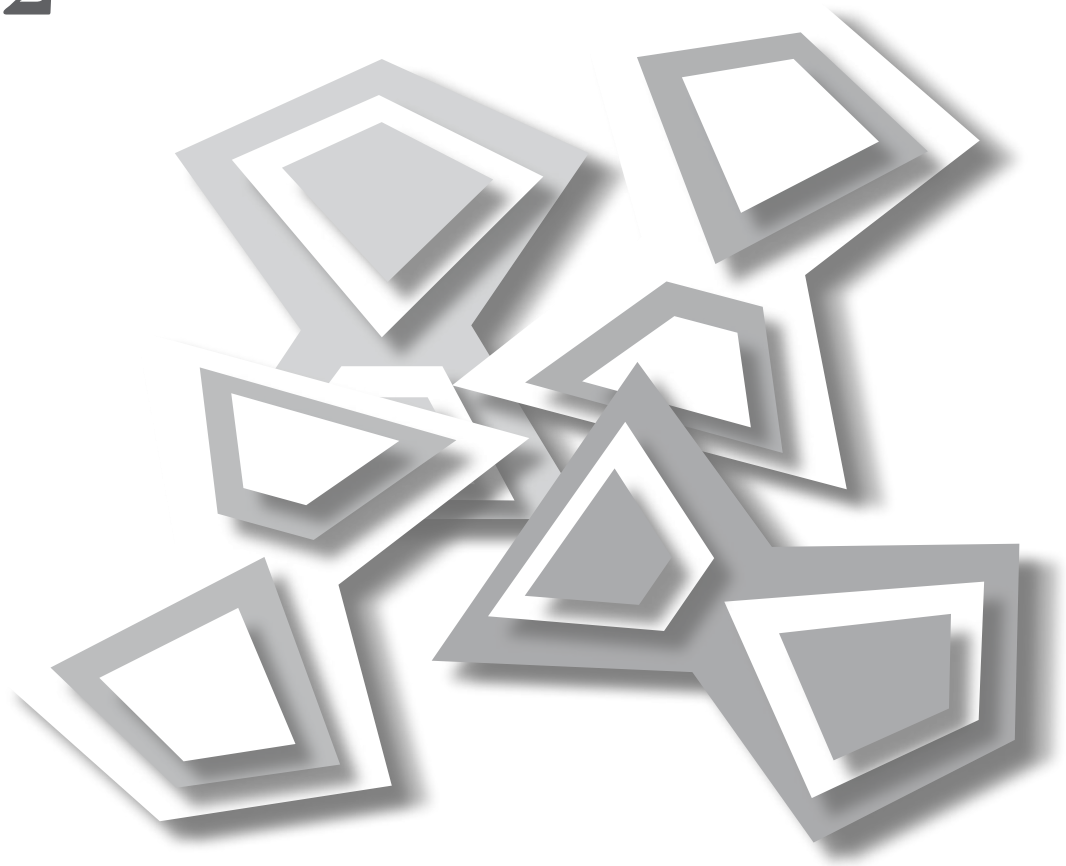
Brian English, "Conducting Ethical Evaluations with Disadvantaged and Minority Target Groups," *Evaluation Practice*, Volume 18, Number 1, 1997

Jody L. Fitzpatrick and Michael Morris (eds.), "Current and Emerging Ethical Issues in Evaluation," *New Directions for Evaluation*, Number 82, Summer 1999

Ethics, Canadian Evaluation Society http://evaluationcanada.ca

Guiding Principles, American Evaluation Association http://www.eval.org/Guiding%20Principles.htm

Chapter **12**



METHODS

This chapter contains:

1. Overview of basic concepts
2. Standard data collection methods
3. Cautionary note about data collection
4. Guiding questions to aid in selecting methods
5. Instrument development and testing
6. Disaggregated data
7. Data analysis
8. Unique peacebuilding tools
9. Sound basis for generalization
10. Record maintenance systems

INTRODUCTION

"All men by nature desire knowledge."

-ARISTOTLE

This chapter introduces peacebuilding practitioners to the options and considerations for selecting the means of data collection for baselines, monitoring, and evaluation. In logical frameworks, methods are referred to as means of verification (MOV), while in other circles they are called research instruments. In this chapter, the term "research" will be used to mean any data collection done for baselines, monitoring or evaluation.

Data collection methods have been developed over a period of many decades and now have well-established standards and techniques.

Data collection methods have been developed over a period of many decades and now have well-established standards and techniques. A comprehensive introduction to methods requires a manual or two of its own. For the purposes of the practitioner, one needs to understand the core concepts and terminology as well as when to use which method for the best results. The techniques of designing and implementing those methods are beyond the immediate needs of the average practitioner, hence the scope of this manual.

Fundamental to research is the notion that data and conclusions are two different things.

It should be made clear what this chapter is not. This chapter is not trying to make researchers out of practitioners. It does not have enough information on how to select, design, and implement methods nor does it cover how to analyze the resulting data such that a beginner could do so effectively. Instead, this chapter is intended to prepare a practitioner to have knowledgeable conversations with professional evaluators in order to make informed choices.

What are the basic concepts I need to know about data collection?

Mastering a few core concepts and their associated terminology is the first step in understanding data collection.

Key Terms

Fundamental to research is the notion that data and conclusions are two different things. Data is the building block of information and is often thought of as statistical or quantitative, although it may take many other forms, such as transcripts of interviews, maps, photographs or videotapes of social interactions. Conclusions are drawn from data through analysis.

Methods are the means of acquiring the necessary data.

Methods are the means of acquiring the necessary data. Once gathered, the data is analyzed in order to generate conclusions.

There are a number of other terms used in research that are important to understand. These terms are listed below in the Methods Key Terminology table, and their accompanying definitions have been made informal and simple.

Methods Key Terminology

Bias	To be inclined toward a particular way of looking at or understanding something
Causality	The direct effect of one event on a future event
Correlation	The extent to which two or more things are related to one another
Disaggregate	To separate into component parts
Generalization	The extent to which one can come to broad conclusions about a group or phenomenon based on information gathered from a set of representatives of that group or phenomenon
Reliability	Do repeated applications of the method, even when different people apply it, result in the same outcome?
Sample	Representative members of the entire client base for the activity
Statistically Significant	Meaningful, measurable relationship or level of change
Unit of Analysis	The primary entity under evaluation, such as individuals, groups, artifacts, geographic units or social interactions
Validity	Does the method measure what it is supposed to?

Quantitative & Qualitative: Methods and Data

Quantitative methods are used to gather data to be analysed in numerical form.

Quantitative and qualitative are important concepts for practitioners to understand. Quantitative methods are used to gather data to be analysed in numerical form. They pose questions of who, what, when, where, how much, how many, and how, generally in the form of surveys and questionnaires. These methods are designed to produce data that tells us how many people do or think something, and which is statistically reliable. Quantitative data typically is in numerical form such as averages, ratios or ranges.

Qualitative methods have greater flexibility and pose questions in a more open-ended manner. They give an in-depth understanding of why people hold particular views. They also explore how people make judgments, in a way that structured quantitative research cannot. Qualitative

methods are not intended to be statistically reliable, but findings can — if participants (those who provide data to the study) are broadly representative — be strongly indicative of the population as a whole. Standard qualitative methods include interviews and focus groups. Qualitative data is typically words or text, though it can include photographs, video, or sound recordings.

Quantitative methods are used to gather data to be analysed in numerical form.

The descriptions provided thus far cover quantitative and qualitative data and data collection methods. It is important to understand, however, that there are also quantitative and qualitative analysis approaches. See the question about analysis on page 50 for more information.

There has been an ongoing debate about which method – quantitative or qualitative – is better suited for baseline, monitoring, and evaluation purposes. Advocates for a quantitative approach argue that their data is hard, rigorous, credible, and scientific. On the other hand, proponents of the qualitative method contend that their data is sensitive, nuanced, detailed, and contextual. For peacebuilding baseline, monitoring, and evaluation purposes, this debate is needless because both approaches are necessary.

Mixed methods – the use of both qualitative and quantitative approaches – are becoming the new norm because they produce a richer set of information to meet the needs of conflict transformation projects. Numbers alone rarely answer the questions that peacebuilders have regarding why and how social change occurred. Perceptions and feelings that cannot be generalized to a greater population do not provide the complete picture either, necessitating the use of quantitative methods.

Qualitative methods have greater flexibility and pose questions in a more open-ended manner. They give an in-depth understanding of why people hold particular views.

What are the standard data collection methods?

All of the standard social science methods, such as surveys or interviews, as well as participatory techniques like mapping, can be utilized in baseline, monitoring, and evaluation of conflict transformation. Each of these methods can be implemented individually or in combination with each other depending on the research needs. This chapter provides short overviews of some of the more commonly utilized methods and their strengths and weaknesses. How to choose the right method for research follows.

All of the standard social science methods can be utilized in baseline, monitoring, and evaluation of conflict transformation.

- **DIRECT OBSERVATION**: Watching, taking notes, and recording specific actions within a target community, such as communi--cations, spatial interaction, or exclusion. The observation can be focused on a project process in which the people participating in the intervention are observed. It can also be focused on changes, such as in people's behaviors and attitudes, which involves watching people go about their daily business at home, in the community, or in the fields.

- **INTERVIEWS**: One-on-one contact with stakeholders, either in person or by telephone. These can be formal structured exercises, where a strict interview protocol is followed, or semi-structured meetings that are partially structured by a flexible interview guide. For comparability purposes, a minimum degree of commonality must exist in unstructured interviews.

- **FOCUS GROUPS**: Small-group conversations that seek to understand how people feel or think about an issue, product, service or idea. Focus groups have a specific purpose, size, composition, and process. They are best conducted with 6-8 people who are selected because they have something in common. Leading focus groups requires a skilled moderator and is best done in a comfortable, permissive environment. Such groups are a compromise between participant observation and more in-depth interviews.

- **PARTICIPANT DIARIES**: These are narrative descriptions of a personal experience. They may be open-ended to allow individuals to capture what was of importance to them each day or week, for example. Participant diaries can also be structured so that individuals take note of specific attitudes, events, behaviors that they have experienced in the allocated timeframe (e.g., daily, weekly).

- **PHOTOGRAPHY/VIDEO**: Utilizing photographs or video to collect visually represented information

- **PROJECT DOCUMENT REVIEW**: Collecting, reading, collating, and analyzing key documents such as proposals, donor reports, annual reports, case studies, etc.

- **QUESTIONNAIRE**: A set of specific, targeted questions to which stakeholders respond in writing. The questions must reflect cultural awareness and be language sensitive in addition to fitting within a set of formal methodological standards. Questionnaires can be distributed electronically, by post or by hand. For a statistically significant conclusion, the number of responses needed depends on the total population size. However, to do statistical applications like developing the mean or plotting charts, one needs a minimum of 30 responses for answers to be valid.

- **SECONDARY DATA REVIEW**: An examination of existing data. This type of review is often the initial inquiry that precedes data collection with stakeholders. It is also called a desk review. Sources include academic theses, annual reports, independent studies by NGOs or researchers, and census data.

- **SURVEY**: A sequence of focused, targeted questions posed to stakeholders in a fixed order by a surveyor. As with questionnaires, survey questions must reflect cultural awareness and be

language sensitive in addition to fitting within a set of formal methodological standards. Surveys are generally utilized for large-scale efforts, though they may be used on a smaller scale. To reach a statistically significant conclusion, the number of responses needed depends on the total population size. However, statistical applications like developing the mean or plotting charts one require a minimum of 30 responses.

- **TESTING**: This is usually a series of questions or exercises – oral or written – for measuring the skills, knowledge, capacities, or aptitudes of an individual or group. Testing is generally used before and after training as a way to measure change.

- **PARTICIPATORY LEARNING AND ACTION TECHNIQUES (PLA)**[27]: The application of Participatory Rural Appraisal techniques in a mutual learning process utilized on broader issues than the original rural development focus. There are many techniques within PLA that are useful for baseline, monitoring, and evaluation purposes such as:

    - **Venn diagrams**: These are made with circular cards of different sizes and colors placed in relation to one another with each card representing an issue. The size of the card represents the issue's importance to the conflict, with a larger card indicating greater importance, and the degree of overlap between cards represents the intensity of interaction of those issues. Men and women, wealthy and poor, young and old, may well produce different diagrams whose differences are often instructive.[28]

    - **Pairwise ranking**: This technique helps to determine the relative importance of various options. The participants compare only two options at a time, and the reasons for preferring one option over the other are made clear. Going through all the possible combinations finally results in a list of criteria by which villagers can assess options.

    - **Conflict Mapping**: This is a technique used to represent a conflict graphically by placing the parties in relation both to the problem and to each other. When people with different viewpoints map their situation together, they learn about each other's experience and perceptions, and their differences and commonalities become clear.

    - **Drawing**: This technique is often called mapping as well. It is a visual depiction, generally in pictorial form, of the focus of the discussion. This can be a geographic map, an emotion, or a situation.

[27] Adapted from Participatory Rural Appraisal (PRA), part of the "Working instruments for Planning, Evaluation, Monitoring and Transference into Action" series(Swiss Agency for Development and Cooperation[SDC], Strategic Unit, January 1997)

[28] Adapted from Participatory Research for Sustainable Livelihoods: A Guide for Field Projects on Adaptive Strategies. http://www.iisd.org/casl/CASLGuide/RepRel.htm.

- **Role Playing**: Taking on a role enables people to creatively remove themselves from their usual roles and perspectives to portray a situation.

Since every context and application of a method can vary, it is difficult to give definitive guidance on what method to use in every situation. Some of the more common strengths and weaknesses of each method are described in the table below. Because cost is almost always a variable in selecting methods, a separate column has been added to illustrate if the method is of high, low, or average expense. Of course, this is also a question of scale: the larger the scale, the higher the cost. Thus, the cost column depicts comparative costs by assuming the same project scale is used in each method listed.

Selecting the Most Appropriate Methods

Method	Strengths	Weaknesses	Cost
Direct Observation	• Minimal preparation required • May enable the experience of minorities or women to be captured, particularly in situations where speaking out against the norm is dangerous	• Must be done at the right moment in the right place • Does not provide information on why things occur • Presence of the observer may influence behaviors	Low
Focus Groups	• Can identify issues that need probing through another method • Allows one to observe various perceptions on an issue • Enables more people to be involved in less time then individual interviews	• Difficult to manage multiple opinions • "Group think" may occur • Individuals may not feel comfortable to dissent	Low/ Average
Interviews	• Good for small numbers • Allows for exploration into how and why • Generates data on needs, expectations, attitudes, perceptions, beliefs, and feelings	• Time-consuming • May be difficult to differentiate between those who are telling you what they think you wish to hear from those telling the truth	Low/ Average

Selecting the Most Appropriate Methods

Method	Strengths	Weaknesses	Cost
Participatory Learning and Action Techniques	• Can offset the biases of the evaluator • Empowers the local people because their views are taken seriously • Excellent methods for working with illiterate communities • Useful if the purpose is to determine whether needs are being addressed by the evaluation (Appropriateness)	• No hard quantified data is produced • For comparative work replication is difficult as each situation has its own unique situation	Low/ Average
Photography/ Video	• Less open to interpretation • Easily disseminated • Can be a rapid technique • Easily led and completed by participants • Easily preserved	• Must be done at the right moment in the right place • Can be one-dimensional information that does not explain how or why	Low/ Average
Project Document Review	• A low-cost way of learning the history and background of a project • Provides insight into the perceptions of the practitioners	• May be limited in the degree of detail • May be tailored to donor or other needs and requirements but omit key information • May have gaps in time that reports do not cover	Low
Questionnaire (by post or e-mail)	• Good if the intervention affects large numbers of people • Good if statistically significant results are needed	• Requires literacy • Time-consuming • Is difficult to utilize in contexts with multiple languages • Requires a distribution system (e.g., postal system, Internet) for large numbers of people	High
Secondary data review	• Fast means of gathering key background information • Offers a variety of perspectives and insights • Can save the evaluation team time since they do not need to collect the data	• May not be tailored to the needs of the project • Data may be flawed	Low

Selecting the Most Appropriate Methods

Method	Strengths	Weaknesses	Cost
Surveys	• Good for interventions that affect large numbers • No literacy requirements for respondents • Good if statistical comparisons are required	• If not well trained, surveyor bias may affect responses • Requires greater resources than a questionnaire does • Does not explain how or why something happened • Not very good if the purpose of the assessment is to study complex processes • Time-consuming	High
Tests	• Good for knowledge acquisition; less reliable for skills	• Does not reveal whether or not the new knowledge or skills will be retained or applied in the future • May require literacy	Low

Data collection is based on standards of practice that are essential to the quality and, hence, the credibility of the data collected.

Data collection is based on standards of practice that are essential to the quality and, hence, the credibility of the data collected. The manner in which questions are asked and formulated, the behavior of interviewers, and identity of the interviewer (in terms of gender, nationality, race, etc.) can influence responses. The most common error made by novices is to allow their personal bias to influence the situation. The type of questions researchers ask can introduce bias as can the choice of who they talk to and when data collection is conducted. In addition, the way that data is analyzed or presented can introduce bias.

When considering how bias can affect the types and form of questions, consider, for example, a project working with French-speaking citizens of Quebec on changing their perspective from that of seeking political separation to remaining part of Canada. One of the attitudes to be changed is animosity toward the federal government. If the baseline asked, "How much do you dislike the federal government?", there is clearly a negative bias that assumes all francophones dislike the federal government and it is simply a matter of how much. A very different response would be expected if the question were phrased as follows: "Tell me about Quebec's relationship with the federal government."

The manner in which questions are asked can also affect the answers given. This becomes particularly important when the evaluator and those giving information are from different cultures. Moving to the other end of Canada, the Blackfoot nation is one of the First Nations (indigenous populations) of Canada living in Alberta. If an evaluator were collecting data from the Blackfoot people, she/he would need to understand their use of silence and not rush to fill it with more questions or answers. The

evaluator who does not allow for silence and the processes of reflection within that community will return with data that is superficial at best and probably significantly flawed.

Finally, who asks the questions also influences responses. The attributes of the questioner are all noticed by the respondent and subtly interplay with the answers given. Formal education, use of special terminology, clothing, being in a paid position, being perceived to be in a position of authority, and being perceived to be sympathetic to the other side due to one's nationality are all examples of the things that may affect the responses received. Consider an older, well-dressed, male French professor going into the immigrant communities of the Paris suburbs following the November 2005 riots to gather baseline data for a project involving disaffected immigrant youth who participated in the violence. The differences between the data collector and the respondents are not fatal to the project; those differences simply need to be taken into account when designing the data collection process and designating who should collect the information.

Ways to minimize bias and errors in data collection include the careful training of researchers, setting of objectives and indicators, and the triangulation of information.

Ways to minimize bias and errors in data collection include the careful training of researchers, setting of objectives and indicators, and the triangulation of information. Because the field of social science research is well-established, there are many ways to obtain training in methods implementation ranging from university degrees in the subject to short courses and practical trainings.

How are methods selected?

Having a clear understanding of what the research exercise (e.g., baseline, monitoring or evaluation) is to explore how the resulting conclusions are to be utilized is essential to good method selection. The information required should drive the selection of methods. Use the following questions to further guide the method choice.

- **HOW COMPLEX IS THE PROJECT?** If the project is extremely complex, qualitative methods are likely to be better suited to handling its intricacies than other methods. As project complexity increases, so too does the need for triangulation.

- **WHAT IS ALREADY KNOWN ABOUT THE PROJECT?** The more data available, the more the evaluators can focus on gathering information to fill the gaps. This may be exploring the why behind a fact in which case qualitative means. On the other hand, it could be that generalized conclusions are missing; hence, quantitative means would be better suited.

The information required should drive the selection of methods.

Determining if a certain design tool was used to plan the project and whether or not the project was launched with clear assumptions and an articulated theory of change provides a sound starting point for an evaluation team. If this information is not available, time should be allotted to finding it through project document review and possibly participatory learning and action techniques with the project team.

- **WHAT INFORMATION IS NEEDED?** Does the team need information on attitudes, knowledge, perceptions, behaviors or skills? If data on knowledge and skills is required, testing might be a good approach. If behaviors are the focus, direct observation or interviews may be a good choice. If it is important for the answers to be generalized, quantitative methods are preferable.

- **HOW MUCH TIME IS AVAILABLE?** Some methods require more time to design, implement, and analyze than others. A large questionnaire, for example, takes far more time than do focus groups. A questionnaire needs to be developed, tested, redesigned, and distributed to respondents, who must be given time to fill it in. Once the questionnaires are returned, the data is generally entered into a database and then analyzed. All of these tasks not only require enough days for each to be performed, but also a sufficiently lengthy evaluation duration to allow for the data collection method to be completed. For information on time and duration estimates, see page 138 of the Evaluation Management chapter. An evaluation for which data collection must be completed in two weeks would not permit enough time for the development of a questionnaire.

- **WHAT IS THE EVALUATION APPROACH?** Although most methods can be applied within most approaches, the approach selected may lend itself better to one method over another. If the approach is self-evaluation and no one on the team has survey or questionnaire expertise, that is probably not the best option. If utilization-focused evaluation is the approach and the project team wants results that can be generalized, surveys or questionnaires would be a good method choice. However, the use of a different utilization-

focused evaluation may be desirable in order to explore why some people drop out of a program and this would best be accomplished through interviews or focus groups.

- **WHAT IS THE EVALUATION SCOPE?** The scope of the evaluation is an important factor in methods selection. If the evaluation is to cover a large number of people, questionnaires and surveys should be considered. For a medium-sized group, participatory learning and action techniques may be a good choice. If the geographic scope is large or the conclusions need to be generalized, time-intensive methods such as interviewing and direct observation would not be the best options.

- **HOW DIFFICULT WILL IT BE TO ACCESS THE DATA?** Also called data availability, the degree of difficulty involved in accessing information is always a consideration in method selection. Are the data sources out in the bush fighting a war and therefore hard to access or are they based in the city centre and accessible through the Internet? The latter may offer the opportunity to use an Internet-based questionnaire tool, while evaluation of the former would require direct observation, interviews or potentially focus groups. Do those you seek to engage speak many different languages? If so, questionnaires and surveys require translation and testing in each language, which will have cost and time implications.

- **WHEN IS THE INFORMATION NEEDED?** If there is a restricted timeframe for the evaluation, rapid methods like secondary data review, project document review or Participatory Learning and Action Techniques may be helpful. If a longer time period is available, more time-intensive methods can be used, such as surveys.

- **WHAT LEVEL OF RELIABILITY IS REQUIRED?** If a high degree of reliability is essential, Participatory Learning and Action Techniques are often not the best options, whereas questionnaires and surveys may be a better choice. The importance of method reliability increases with the emphasis on quantitative results and the ability to generalize to populations.

- **WHAT IS THE AVAILABLE BUDGET?** As outlined in the Selecting the Most Appropriate Methods table on page 207, different methods have different cost structures associated with them. Direct observation and project document review are very low in cost, while questionnaires and surveys can be expensive.

- **WHAT IS THE CAPACITY OF THE DATA COLLECTORS?** The difficulty of the method should be considered in relation to the capacity of the data collectors. If project staff members are gathering the data, selection of a simple and easily applied method is important to the quality of the data collected. Try to build on the skills

that the staff already have acquired through their project implementation work by choosing, for example, such methods as conflict mapping or photography/video. In some cases, the method may appear simple, but actually gathering data that contributes to sound decisionmaking may be more complex. The greater the degree of complexity required by the method (also called the method difficulty), the greater the level of knowledge and skills required to implement it effectively.

- **WOULD TRIANGULATION MAKE THE CONCLUSIONS MORE RELIABLE?** Triangulation is simply using different methods to research the same issue and then analyzing all of the results. For instance, if examining the results of a project that sought to change laws on citizenship in Ivory Coast, one could conduct in-depth interviews with members of the government, use conflict mapping with groups of non-citizens in their communities, and gather feedback forms from events. Here, the evaluation is cross-checking one result against another (i.e., triangulating), which increases the reliability of the conclusions.

Triangulation is simply using different methods to research the same issue and then analyzing all of the results.

Triangulation is useful in many ways. Contradictory results produced through different methods often indicate important problems with question design and/or fundamental issues surrounding the researcher's understanding of a topic. Triangulation is essential when using Participatory Action and Learning Techniques and helpful when the researcher, in exploring sensitive issues, is uncertain if the data source is able or willing to provide the full story.

The conflict and its volatility affect the choice of method.

- **ARE THERE CONTEXTUAL ISSUES SUCH AS CULTURE OR LEVEL OF DEVELOPMENT THAT WOULD MAKE SOME METHODS BETTER THEN OTHERS?** The context should always be considered in method selection. If women will not speak openly in front of men or youth cannot speak their mind in front of elders, then focus groups or surveys conducted in open environments may not be the right choice. Individual interviews or questionnaires may allow those who cannot speak their mind in all situations better opportunity to express their views.

- **ARE THERE CONFLICT ISSUES THAT MAKE SOME METHODS BETTER THAN OTHERS?** The conflict and its volatility affect the choice of method. There may be conflict situations where someone is unable to state a dissenting opinion without putting themselves in danger. In such a case, methods should be selected that ensure the anonymity of sources by not requiring the disclosure of names or documentation. Any method that requires experiences to be documented, through the use of participant diaries or photographs, for example, deserves extra consideration in conflict contexts. If discovered by the wrong people, such as a paramilitary group or the army, these participants and sources might be in danger.

Take, for instance, three selection criteria mentioned earlier – data availability, method difficulty and method reliability – and compare their implementation in stable versus volatile environments. Recall the definitions for each:

- **Data availability** reflects the ease or difficulty of obtaining data. High data availability means that it is relatively easy to access the data from the source. Low data availability would mean that is difficult to obtain the information needed.

- **Method difficulty** considers the complexity of appropriately developing and implementing the method. High method difficulty implies that there is some intricacy involved in developing and implementing the instrument, while low method difficulty means that it is relatively simple.

- **Method reliability** refers to the degree to which the method produces the same results when used by different people. High method reliability means that the instrument can be used many times and the same responses will be generated. Low reliability means that, if different people utilized the method, the answers would likely be different.

  In the table on page 215, the left half of each column represents a stable environment in which a non-sensitive issue is discussed. The right half of each column represents a highly contentious conflict situation dealing with a sensitive issue.

  For example, consider a project in the Ukraine that seeks to decrease negative attitudes of Ukrainians toward Russian nationals living in the Ukraine. Using focus groups would provide high data availability since accessing this information would not be difficult. In comparison, consider a similar project in Iraq in 2005 that seeks to decrease negative attitudes that Sunnis hold toward Americans. The use of focus groups would likely have low data availability because Sunni respondents would be fearful of speaking out against the accepted norm of '"hatred" in front of their community.

Data Availability, Method Difficulty and Reliability[29]

Methodology	Data Availability		Method Difficulty		Method Reliability	
	Stable	*Volatile*	*Stable*	*Volatile*	*Stable*	*Volatile*
Direct Observation	High	High	Low	Low	High	High
Focus Group	Average	Low	Average	High	Average	Average
Interview	Average	Average	Average	High	Average	Average
Participatory Learning & Action Techniques	Average	Low	Average	Average	Low	Low
Photography/Video	High	High	Low	Average	High	High
Project Document Review	High	High	Low	Low	High	High
Secondary Data Review	High	Low	Low	Low	High	High
Survey/Questionnaires	Low	Low	High	High	High	Average
Tests	Average	Average	Average	Average	Average	Average

What do I need to know about instrument development and testing?

Once the methods have been selected, the next step is to design the instrument to be used, such as the questionnaire or the interview protocol. The methodological standards for instrument development are based on a well-researched body of literature, and it is critical to carefully follow those standards. The result of not using these standards will be flawed or biased instruments that produce unreliable data. That flawed data will then create false conclusions. Since the conclusions of an evaluation are what inform program decisionmaking, the consequences of using flawed instruments can have significant negative effects on the project and the people it is meant to assist.

At this point in the process, the average practitioner should seek technical assistance if developing the instrument on her/his own or assign it as one of the tasks for the external consultant.

Once the instruments have been designed, they must be tested. Both qualitative and quantitative methods of data collection require validity and reliability tests. These check for clarity, accuracy and whether the

[29] Adapted from Managing for Evaluations: A Manual for PVOs. PVC Grants Managers and Evaluation Team Members by A. Willard (USAID/PVC/Matching Grants, 1998).

tool is appropriate to the data needs. In essence, you want to know whether or not the instrument accesses the type of data required in the way that is intended. This step is commonly omitted in peacebuilding evaluation to the detriment of the quality of results. When external consultants are hired, they should be told that they will be expected to perform validity and reliability testing and this expectation should be included in the terms of reference (see page 138 for more information on terms of reference development.) If instruments are not tested and refined, the quality of the resulting data will be compromised.

Once the instruments have been designed, they must be tested.

Testing can be as elaborate or as minimal as is needed by the project at hand. Assuming that your project staff members are broadly representative of both the local population and the conflict, a cheap and easy strategy is to use the project staff as subjects for the test. In other words, try the instrument on project staff, by having them fill out the questionnaire or respond to an interview. This will not only enable the evaluators to have input into refining their instrument, but it will also provide the project team with insight into what is being asked and how. Of course, if the instruments are intended to gather data from Israelis and Palestinians but the staff members of the evaluating organization are all Israeli, this would not be an effective test. Similarly, if a consensus-building project in Washington, DC worked primarily with community members who have limited educational backgrounds, testing the instruments on highly educated staff would not be appropriate. Always test the instruments in the language and culture in which they will be used.

Testing can be as elaborate or as minimal as is needed by the project at hand.

What is disaggregated data?

In thinking about the data to be sought from the baseline, monitoring or evaluation, consider whether the conclusions would be more useful if they were broken down according to different factors such as ethnicity. These factors correspond to key groupings within the project's target population. This is called disaggregated data.

It is also important to consider the state of the conflict and the sensitivity of the factor when making these decisions.

Take, for instance, a dialogue project in Northern Ireland that brings together Protestants and Catholics. Attitudinal information that represents all participants gathered through interviews will be informative, but it could hide important differences between the two communities. Potentially, if the data was broken out by Catholic participants and Protestant participants, it could reveal differences in attitudes that would be key to informing future decisions about the project.

Data can be disaggregated by many different factors or subgroups. The importance of each factor is dependent upon the focus of the intervention being explored and the evaluation objectives and lines of inquiry being chosen. Disaggregating the data within a line of inquiry should provide more useful information than if it were not disaggregated.

Data can be disaggregated by many different factors or subgroups.

If the information is not more useful when disaggregated, it is not worth doing it. Common factors to disaggregate against for conflict transformation projects include:

- Gender
- Residence
- Ethnicity
- Religion
- Age
- Tribe
- Ex-combatant
- IDP or Refugee

Remember that, as the number of factors increases, so too does the amount of data needed and the amount of time required to analyze it, which affects the cost of the research.

Like instrument development and implementation, data analysis is subject to strict standards of practice.

It is also important to consider the state of the conflict and the sensitivity of the factor when making these decisions. If asking someone about her/his religion or ethnicity is highly sensitive or even, in some instances, dangerous, there might be proxies that can be used to substitute for such questions as a way to avoid endangering the respondents. For instance, in some conflicts, neighborhoods are strictly segregated, so that asking where someone lives could substitute for ethnicity or tribe. The project team plays a key role in alerting external professionals to these sensitivities and to potential proxies.

Quantitative analysis can also be used on qualitative data.

Once the factors have been chosen, the appropriate questions need to be included in the research instrument in order to gather the data. For instance, if the gender and religion of participants were important to the project, these two questions would be added to a questionnaire or interview as factors to be collated. Or, if a feedback form is used to monitor a workshop, these two factors (gender and religion) would be added to the form.

How is data analysis done?

Like instrument development and implementation, data analysis is subject to strict standards of practice. Even though high-quality data collection instruments have been developed and tested appropriately, and the process of data collection may have followed all the best practices, it all will be for naught if the proper analysis techniques are not utilized.

Quantitative data analysis is usually called statistics. Generally speaking, quantitative data is processed through statistical computer software packages. The most popular of these is the Statistical Package for Social Sciences (SPSS); however, the analysis software could be as simple as an Excel spreadsheet or an Access database. With the proper safeguards for confidentiality, processing quantitative data in a public or transparent way helps build credibility in communities where people distrust the process and/or technology.

There are a number of approaches to qualitative data analysis, although none have the same precision in the rules as statistical analysis. The abundance of approaches is an asset because any set of qualitative data can be approached from a variety of different perspectives. In other words, different techniques can be applied to the same data, which may highlight new aspects of that information. Hence, the approach taken in the analysis affects the conclusions drawn.

There are a number of approaches to qualitative data analysis, although none have the same precision in the rules as statistical analysis.

Quantitative analysis can also be used on qualitative data. This means that qualitative data, such as transcripts of interviews, can be analyzed to produce statistical conclusions.

Identifying the means of analysis is often forgotten when discussing baselines, monitoring, and evaluation, yet it is a central part. Peacebuilding practitioners should question evaluators about their chosen means of analysis and why they believe it is the best option. For all forms of research, the author needs to be able to explain how she/he arrived at the conclusions from the data. In other words, the means of analysis must be described and scrutinized.

Analysis should not be the exclusive domain of the evaluator. The people under scrutiny in an evaluation frequently offer very insightful analysis and, at times, bring out dimensions that only they can perceive. Whoever participates in the analysis needs to be aware of her/his own biases and assumptions.

When a standard data collection method is refined or blended with core features of a field it is called a tool.

Are there unique peacebuilding tools for data collection?

In many professional fields, such as education, basic social science data collection methods are blended together or refined to meet specific needs of that field. When a standard data collection method is refined or blended with core features of a field it is called a tool. These '"tools" often become the accepted techniques for baseline, monitoring, and evaluation in that field.

Few tools have been developed specifically for conflict transformation, although there are some that were developed for other fields that could be adapted. None of the tools in this latter group have taken root in the peacebuilding field, none have been deemed more or less effective than others, nor is there a norm or standard in terms of application.

Selecting a tool is not essential for an evaluation to be effective. Evaluators and project teams should consider what information is being sought and then select the best way of obtaining it. That selection may include the use of straightforward data collection methods, such as surveys or interviews, or there may be another tool that is more effective in obtaining the data.

Each tool needs to be adapted to the context and purpose of the intervention for which it is being used to evaluate and, as with methods, tools can be combined. Regardless of the tool used, it is necessary to have an understanding and ability to implement the data collection methods from which they are developed.

As of 2005, the Office of Transitional Initiatives (OTI/USAID) is developing a monitoring and evaluation toolkit for transition and conflict transformation projects.[30] A sample of the tools in this toolkit, along with several others, is listed below. Note that this is not a comprehensive list of every tool that has been developed – others are available.

ACTIVITY INTERVIEW[31]: An activity interview seeks to identify the views participants hold about an intervention's activity, such as a Good Neighborliness Seminar, regarding the effects of that activity and/or to obtain process feedback. This input is added to the observations and opinions of the staff responsible and compiled in a short report. Using a semi-structured format, activity interviews take place a few weeks after the activity. Not all participants need to be interviewed, nor does it need to occur after each activity; rather, a sample of each may be selected. This tool would most commonly be used in monitoring.

COGNITIVE SOCIAL CAPITAL ASSESSMENT TOOL (CSCA): CSCA is "a quantitative method for collecting basic information about cognitive social capital quickly. Cognitive social capital refers to people's perceptions of the trustworthiness of other people and key institutions that shape their lives, as well as the norms of cooperation and reciprocity that underlie attempts to work together to solve problems."[32] CSCA uses a questionnaire, which can also be utilized as a survey. It can be implemented in small- to large-scale applications. The quantitative nature of the tool may fail to capture some of the complexities of social change. This tool is most appropriate for a baseline and evaluation.

MEDIA CONTENT ANALYSIS TOOL: This tool allows project managers to "evaluate media coverage, placement of stories, tone, and visual images, prominence of quotes/personalization, and reach of a media outlet. It can be used to track how different media cover topics such as conflict, human rights, reintegration of ex-combatants, and local governance reform."[33]

To utilize this tool, the evaluator identifies the media outlets to include, determines their circulation/reach, and then scores each against seven measures: prominence, headline, visuals, quotes, tone, column inches, and political ideology. The score for each outlet is then multiplied by

[30] At the time of this manual's publication, the *OTI Performance Monitoring and Evaluation Toolkit, Fast Learning for Program Improvement, Local Empowerment and Results [*hereinafter *OTI Toolkit]*, is not yet available publicly. Inquiries should be made to Social Impact at http://www.socialimpact.com/.

[31] Adapted from *Strategic and Responsive Evaluation of Peacebuilding: Towards a Learning Model, Report of the Second Action-Reflection Seminar Convened by NPI-Africa and the NCCK-CPBD Project*, March 2001.

[32] *OTI Toolkit*, forthcoming, pp. 146.

[33] *OTI Toolkit*, forthcoming, pp. 111.

the outlet's individual ranking in terms of circulation/reach. The scoring measures may need to be adapted to those issues of importance to conflict transformation. This tool is best used for baseline and evaluation studies, though it could be modified to contribute to monitoring efforts.

CASE STUDY: This tool investigates a contemporary event within its real-life context. Case studies are a way to learn from past experience since they explore how something happened. A case study results in a report that contains a rich narrative of the phenomenon detailing how it came about.[34] It is based on a particular worksite the boundaries of which need to be clearly defined to allow the study to be focused. If the case is to be illustrative of the wider context, selection of a site that is not unique is important.

Case studies rely on multiple data sources because the data needs to converge in a triangulating fashion. A case study involves the use of a variety of data collection methods, predominately interviews and project document review, but can also include direct observation or focus groups. Case studies can be done on a rolling basis to monitor the changes that occur over time or as part of an evaluation both of which result in brief, reflective snapshots of complex and dynamic situations.

CAPACITY ENHANCEMENT NEEDS ASSESSMENT (CENA): CENA "is a participatory assessment [tool] designed to evaluate existing capacity within key community and local government stakeholder groups, identify capacity gaps and weaknesses and recommend possible remedies."[35] Based on interviews and focus groups, information from a CENA is then plotted against each indicator on a scorecard. The tool was created for community-based development and, as such, the indicators would need to be adapted for conflict transformation programming. This tool would be best used in baseline and evaluation studies, though a streamlined version might be possible for monitoring efforts.

FOUR LEVELS OF TRAINING EVALUATION[36]: The four levels approach - reaction, learning, transfer and results - is a systematic way to assess the quality and outcome/impact of training. Information from the previous level serves as a base for the next level as one works through all four. Reactions (level one) should never be the only level utilized and can often be blended with learning (level two). Generally, reaction (level one) utilizes a questionnaire and learning (level two) involves a pre- and post-test, while transfer (level three) can either use interviews or surveys several months after the training has taken place. Levels one and two should be included in monitoring systems while levels three and four might be a monitoring or evaluation tool.

[34] Robert Yin, *Case Study Research: Design and Methods*, Third Edition, Thousand Oaks, Sage Publication, 2003.
[35] *OTI Toolkit*, forthcoming, pp. 55.
[36] Adapted from the *OTI Toolkit*, forthcoming.

Can I draw a conclusion for the entire population from this data?

How to draw conclusions for an entire population, be it a community, tribe, or geographic area, is beyond the scope of this introductory manual. Conclusions are drawn using quantitative methods and should be based upon "enough" of a representative or random sample of the population. What constitutes "enough" requires a calculation based on established practices.

As practitioners who will be gathering data through monitoring practices, however, there is one related concept that is important to understand. It can be called informally *Lovers, Haters, and Everybody Else*. As depicted below, the bell curve represents the average population of people, with the majority located in the middle and the extremes located at either end. Extreme means not being representative of the average or norm within the group. Although it appears that the extremes are the same size at either end, this is not necessarily the case.

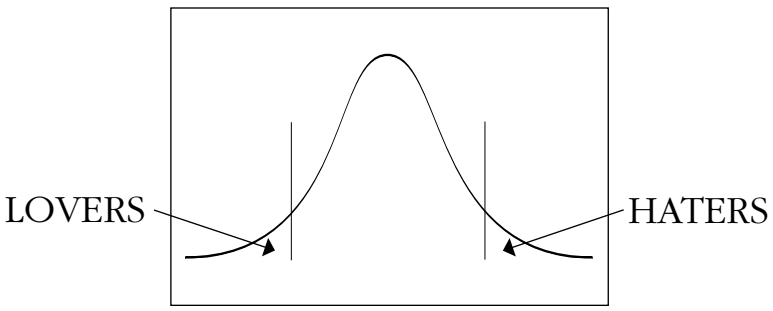
In other words, they are not representative of the average – they are lovers or haters.

Lovers, Haters, and Everybody Else relates to conflict transformation monitoring because it explains the implications of basing our monitoring data collection only on participant-initiated measures such as personal narratives, testimonials, fan mail, individuals who call into radio shows, or informal comments from participants. As a general rule of thumb, individuals who are willing to step forward and to take their own time to volunteer information fall at either end of the bell curve. They either love what is being done and want to praise it so it continues or they hate what is happening and feel compelled to intervene to stop it. In other words, they are not representative of the average – they are lovers or haters.

Classifying such individuals as lovers or haters does not discount their input. What it does tell us is what those two ends of the spectrum think about the work. We cannot take that work to the next step and assume that our information is representative of everyone involved.

.

The Lovers and Haters and Everybody Else



What record maintenance systems are necessary for collected data?

Record maintenance for baseline, monitoring, and evaluation data is quite straight-forward. For baselines and evaluations, the raw data collected, such as completed questionnaires or interview notes, should be kept in its hard-copy form. Future evaluators and researchers may wish to go back to the original data and do a different analysis or verify the previous one. If the data is sensitive or confidential, it should be stored appropriately, either in a locked filing cabinet or an inaccessible room. How long raw data is kept is dependent upon the project, organization, and potential future uses. If a formative evaluation was performed and a summative evaluation will occur later, the records should be saved until the summative evaluation is done.

Future evaluators and researchers may wish to go back to the original data and do a different analysis or verify the previous one.

Evaluators often do not return raw data to the project team unless it is requested. Particularly in the case of baseline and formative evaluation, retaining the raw data should be part of an organization's good practice norms.

For the material itself, one can use well-labeled filing cabinets, computer floppy disks, or CD-ROMs to preserve the information. Generally, it is good practice to keep a back up file of all electronic documents. If the data was processed using Excel spreadsheets or SPSS, keep those electronic files because it will save data entry time for future efforts.

If the data is sensitive or confidential, it should be stored appropriately, either in a locked filing cabinet or an inaccessible room.

When considering record maintenance systems for monitoring data, simple is always better. If most project staff members are comfortable with using an Excel spreadsheet, select that as the medium for data analysis rather then a statistical package that requires a specialist to operate.

What are the ethical obligations of feeding back the results to the people involved?

When people hear nothing about a study or evaluation in which they participated, they are less inclined to contribute to future data collection efforts. This is even more the case in highly participatory exercises where people feel they have an investment in the outcome of the research. Even the briefest feedback on the general findings and the use of those findings is often appreciated very much.

Further Reading:

Chronic Poverty Research Centre Methods Toolbox
http://www.chronicpoverty.org/CPToolbox/toolboxcontents.htm

William M. Trochim, *The Research Methods Knowledge Base*, Atomic Dog Publishing, 2001. http://www.socialresearchmethods.net/kb/

Keith Punch, *Introduction to Social Research Quantitative and Qualitative Approaches*, SAGE Publications, 2000.

CONCLUSION

"It is not that we should simply seek new and better ways for managing society, the economy and the world. The point is that we should fundamentally change the way we behave."

- VACLAV HAVEL

An historian once said, "Most of our so-called reasoning consists in finding arguments for going on believing as we already do."[35] Monitoring and evaluation offer the opportunity to counter this tendency, to open ourselves to new information, and to engage in responsible analysis. To exploit this opportunity, greater effort must be made to incorporate design, monitoring, and evaluation into conflict transformation programs. This effort will serve the peacebuilding field by increasing our understanding of, and ability to prove change in complex conflict situations.

Much has been learned from DM&E in peacebuilding so far…

When Boston wanted to stop youth violence and homicide, a partnership – composed of researchers, community leaders, members of the clergy, probation officers, police officials, and federal enforcement agency personnel – came together to devise a strategy to intervene in the local gun market. When data revealed that the problem was more specifically caused by youth gangs, not simply gun markets, the partnership adjusted its strategies. Boston's hard work paid off: youth homicides fell by two-thirds after the ceasefire strategy was put in place.[36]

In considering the importance of involving local authorities in refugee reintegration projects in Rwanda and Bosnia, an evaluation team concluded: Local and regional authorities were taken into account in different ways. Some were included, some excluded intentionally, and some ignored. Where it was possible to include them constructively, they became important allies in promoting coexistence. Where they were ignored, they undermined success.[37]

More dedicated thinking and testing of DM&E techniques is necessary so that they, too, become increasingly effective for peacebuilding. Though strides have been made, there is still much to learn in order for DM&E to reach its full potential as a learning tool within conflict transformation.

One can learn to build with stone by reading books and experimenting. The principles of laying stone are few and easy to follow – enough

[35] James Harvey Robinson, *The Mind in the Making*, http://en.thinkexist.com/quotes/james_harvey_robinson/.

[36] National Institute of Justice, *1998 Annual Report to Congress*, December 1999.

[37] Eileen F. Babbitt et al., *Imagine Coexistence: Assessing Refugee Reintegration Efforts in Divided Communities*, The Fletcher School of Law and Diplomacy, Tufts University, July 2002.

so that progress comes quickly in making walls both aesthetically appealing and structurally sound. By taking apart a stone house built by experienced masons, the novice stone layer discovers how the veterans dealt with more nuanced and challenging issues. The novice can then recycle those stones and build another house in ways she/he had not known previously.

We've tried to disassemble design, monitoring, and evaluation for peacebuilding by taking each subject apart "stone by stone." Every reader will have to gather the stones presented here, and others from elsewhere, to build their own structure. Hopefully, the construction will include approaches and ideas that our readers had not known or previously practiced.

Of course, many of the building blocks needed to succeed in peacebuilding have not been addressed in this manual. Our intent is not to tell anyone how to transform conflicts, but rather to illustrate the range of choices in peacebuilding that catalyze needed change and the techniques available for that work to contribute to learning within the field.

We've maintained throughout this manual that monitoring and evaluation are the most accessible learning disciplines available to peacebuilding. In putting this manual together, we're reminded of another valuable learning discipline: writing. We've forced ourselves to support our beliefs with examples, to check our jargon, and to think backwards and forwards on how we've done and will do many of the tasks here described. We've done less well at shedding our Western paradigms and values. The acts of writing, explaining, and connecting thought and action have proven to be very educational for the authors, as they can be for all peacebuilders.

In closing, we would like to draw attention one more time to a number of themes that run throughout this manual:

- Peacebuilding and DM&E are all about change.
- The most brilliant and creative ideas only become the best alternatives to bringing about change when they incorporate sound practices in design, monitoring, and evaluation.
- Peacebuilding is a unifying process. Successful design, monitoring, and evaluation bring together the parties, practitioners, designers, and evaluators. The artists and the technicians together can produce quality peacebuilding initiatives.
- We can all make better decisions. The data collected through a baseline, monitoring exercise, or an evaluation should inform our decisions, and better decisionmaking can, in turn, improve our work.

- Knowing when to get help is important. Knowing how to use help is imperative.

- DM&E is an opportunity. Baselines, monitoring, and evaluation provide the peacebuilding field with a valuable opportu nity to show policymakers, the public, the press, and the people who we work for – the stakeholders – that conflict transformation produces positive results.

In closing, the old adage for university professors, "publish or perish," comes to mind. For peacebuilding, there is a new standard rightly being imposed: demonstrate effectiveness or perish. The concepts, tools, and examples in this introductory manual are intended to help peace workers of all stripes rally to this call for results. We believe peacebuilding works and that we have a responsibility to show others the results.

"This is not the end. It is not even the beginning of the end. But it is, perhaps, the end of the beginning."

- WINSTON CHURCHILL

APPENDIX A

SOURCES FOR THE TERMINOLOGY DECODER

Organization	Source	Website
SFCG		http://www.sfcg.org
UNDP	UNDP Programming Manual, January 2003	http://www.undp.org/bdp/pm/chapters/progm4.pdf
CIDA	CIDA The Logical Framework: Making it Results-Oriented December 2002	http://www.acdi-cida.gc.ca/cida_ind.nsf/0/c36ebd571b6fa02985256c620066cd6f?OpenDocument
EuropeAid	EuropeAid Aid Delivery Methods: Project Approach "The Logical Framework Approach" March 2004	http://europa.eu.int/comm/europeaid/qsm/project_en.htm#2.%20Logical%20Framework%20Approach
TACIS/ EuropeAid	EuropeAid Standard Call for Proposal Documents "Annex C: The Logical Framework" 2003	http://europa.eu.int/comm/europeaid/tender/gestion/pg/e03_en.htm
DFID	DFID Tools for Development Section 5.3 September 2002	http://www.dfid.gov.uk/pubs/files/toolsfordevelopment.pdf
World Bank	World Bank Log Frame Handbook January 2005	http://www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/2005/06/07/000160016_20050607122225/Rendered/PDF/31240b0Lfhandbook.pdf
UNHCR	UNHCR Project Planning in UNHCR: Practical Guide March 2001	http://www.the-ecentre.net/resources/e_library/doc/Project%20Planning%20in%20UNHCR.pdf
SIDA	SIDA The Logical Framework Approach 2004	http://www.sida.se/Sida/jsp/polopoly.jsp?d=1265&a=16274
RELEX	RELEX ECHO Manual Project Cycle Management October 2003	http://europa.eu.int/comm/echo/pdf_files/partnership/pcm_echo_en.pdf
SADC	Swiss Agency for Development and Cooperation Setting It Up: Definition Logical Planning Framework (no date)	http://web.mit.edu/urbanupgrading/upgrading/issues-tools/tools/Log-Plan-Framework.html
EIDHR	European Initiative for Democracy and Human Rights	http://europa.eu.int/comm/europeaid/projects/eidhr/cfp-micro_en.htm
Danida	Danida Logical Framework Approach: A Flexible Tool for Participatory Development 1996	http://amg.um.dk/en/menu/TechnicalGuidelines/LogicalFrameworkApproach/
USAID	Building a Results Framework, USAID Center for Development Information and Evaluation	http://pdf.dec.org/pdf_docs/pnaca947.pdf

APPENDIX B

EVALUATION TERMS OF REFERENCE (TOR) CIRCULATION OPTIONS

Evaluation field

1. M & E News http://www.mande.co.uk/cgi-bin/www.mande.co.uk/forum.pl
2. PARC (Performance Assessment Resource Center): Submit inquiry through http://www.parcinfo.org/enquiry.asp?subject=consultancy
3. International and Cross Cultural Evaluation Topical Interest Group of the American Evaluation Association: XCeval@topica.com
4. AEA Job Bank http://www.eval.org/JobBank/jobbank.htm
5. African Evaluation Association: http://www.afrea.org/
6. The Evaluation Center EVALJOBS listserv http://evaluation.wmich.edu/archives/index.html
7. Brazilian Evaluation Network http://www.avaliabrasil.org.br/IntroductionRebramaEnglish.html
8. Canadian Evaluation Society listserv: http://www.evaluationcanada.ca/site.cgi?s=4&ss=6&_lang=an
9. UK Evaluation Society http://www.evaluation.org.uk/
10. European Evaluation Society http://www.europeanevaluation.org/
11. Malaysian Evaluation Society http://www.mes.org.my/
12. Nigeria Monitoring and Evaluation Network http://66.201.99.156/afrea/content/index.cfm?navID=3&itemID=735

More national evaluation societies, which do not have listserv or job board, can be found at http://www.mande.co.uk/societies.htm

Conflict Transformation field

- Alliance for Conflict Transformation

http://conflicttransformation.org/Home/SubscribeTermsofService/PostYourAnnouncement/tabid/175/Default.aspx

- Alliance for Peacebuilding: http://www.aicpr.org/

Development field

- Relief Web http://www.reliefweb.int/vacancies/
- DevNetJobs. http://www.devnetjobs.org/
- ConsultingBase: http://www.consultingbase.com/
- Development Opportunities http://www.dev-zone.org/jobs/