

Selecting Evaluation Approaches

Options in Peacebuilding Evaluation



**Prepared for CARE
International UK**

**by Mark M. Rogers
February 2012**

Table of Contents

I. Factors Influencing the Choice of an Approach	3
A. Clarity of Purpose	3
B. Attention to Process	5
C. Coherence within the Overall Evaluation Plan	5
II. Evaluation Approaches Useful in Peacebuilding	7
A. Process-centred Evaluation Approaches.....	7
<i>Participatory Evaluation</i>	7
<i>Action Evaluation</i>	8
<i>Empowerment Evaluation</i>	10
<i>Utilization- Focused Evaluation</i>	12
<i>Goal-free Evaluation</i>	16
B. Change-centred Approaches to Evaluation.....	16
<i>Outcome Identification/Objectives-based Studies</i>	17
<i>Theory-Based Evaluation</i>	19
<i>Impact Evaluation</i>	22
<i>Developmental Evaluation</i>	24
<i>Thinking about Evaluation Approaches</i>	25
III. Cost Considerations	27
IV. Conclusions	30
V. Resources.....	31

Evaluation Approaches

This paper focuses on the evaluation approaches useful in evaluating peacebuilding and conflict-sensitive development programmes lead by NGOs. The intent is to provide programme managers with information they can use to discuss approach options with evaluators.

There is no single evaluation approach that is best for peacebuilding programmes any more than there is one best approach to evaluating health, for example. Nor is there a “good enough” approach. All the evaluation approaches discussed here work best when developed as early as possible (preferably during the programme design) and when reviewed immediately prior to implementing the evaluation.

I. Factors Influencing the Choice of an Approach

When thinking about or designing an evaluation, the evaluation approach is a critical and early consideration. Selecting evaluation approaches requires clarity of purpose, attention to process, and coherence within the overall evaluation plan. Here evaluation approaches have been separated into process-centred and change-centred approaches. Many evaluations combine two or more approaches, particularly a process-centred approach with a change-centred approach. Cost considerations bear more directly on the data collection and analysis methods, rather than the approach. Most approaches are pursued through a range of data collection and analysis methods.

A. Clarity of Purpose

Selecting an appropriate approach or combination of approaches begins with clarity about the purpose of the evaluation and the types of questions the evaluation is intended to address if not answer. Patton identifies six distinct evaluation purposes, paraphrased below:¹

Accountability – demonstrate that the resources are well-managed and efficiently attain the desired results

Judgement/Valuing of overall value – inform decisions relating to the value and future of the programme

Programme Development – adapt the programme in complex, emergent and dynamic situations

Monitoring – manage the programme, identify problems early

Learning – improve the programme

Knowledge Generation – enhance general understanding and identify generic principles

¹ Patton, Michael Quinn, *Utilization-Focused Evaluation: The New Century Text*, 4th edition. SAGE Publications Inc., 2008

An evaluation may have more than one purpose, yet generally speaking it is difficult to resource more than two purposes within a relatively rapid programme evaluation. Evaluations with multiple purposes should prioritize the purposes so that resources can be allocated accordingly. Each evaluation purpose addresses a different set of evaluation questions. The following table² illustrates the coherence between purpose, key evaluation questions and common evaluation approaches. The approaches in red are those that will be covered in greater detail in this paper. These are the approaches most relevant to peacebuilding evaluation.

Purpose	Lines of Inquiry	Common Evaluation Approaches
Judgement of overall value and support major decision making	Does the programme meet the participants' needs? To what extent does the programme have merit? Can the outcomes be attributed to the intervention? Is the programme theory clear and supported by the findings? Is this an especially effective practice that should be funded and disseminated as a model programme?	Theory-based Evaluation, Impact Evaluation
Learning – improve the programme	What works and doesn't work, for whom in what context? How can results be increased or improved? How can quality be enhanced?	Outcome Identification/Objectives-based studies, Reflective Practice, Appreciative Inquiry, Action Evaluation, Empowerment Evaluation
Accountability – demonstrate that resources are well managed and efficiently attain desired results	Are goals and targets being met? Are indicators showing improvement? Are funds being used for intended purposes? Are staffers qualified? Are only eligible participants being accepted into the program? Are resources being efficiently allocated?	Programme Audits, Performance Measurement and Monitoring, Accreditation and Licensing, End of Project Reports, Scorecards
Monitoring – manage the programme, routine reporting & early problem identification	Are inputs and processes flowing smoothly? Are participation rates changing? Are outputs as anticipated and on schedule? Where are there bottlenecks? What variations are there across subgroups or sites?	Management Information Systems, Quality Control Systems, Performance Appraisals
Programme Development -- Adaptation in a complex, emergent and dynamic situation	How is the intervention connected to and affected by the larger changing system? How do we distinguish signal from noise to determine what to attend to? How do we respond to what we cannot control, predict or measure?	Developmental Evaluation, Systems Evaluations, Real-time Evaluation Environmental Scanning

² Ibid.

<p>Knowledge Generation – Enhance general understanding & identify generic principles about effectiveness</p>	<p>What are the general patterns and principles of effectiveness across projects and sites? What lessons are being learned? What principles can be extracted to inform practice? How do evaluation findings triangulate with research results, social science theory, expert opinion, practitioner wisdom and participant feedback?</p>	<p>Cluster Evaluation, Meta-analysis, Effective Practice Studies</p>
--------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------

B. Attention to Process

Among the approaches covered here, there are four that are cross-cutting – they may be employed regardless of the purpose of the evaluation. For these four approaches, the way in which the evaluation is conducted is as important as the purpose of the evaluation:

- Action Evaluation
- Empowerment Evaluation
- Utilization-Focused Evaluation
- Goal-Free Evaluation

This means that an evaluation for the purpose of accountability, for example, might apply Utilization-Focused Evaluation with Theory-based Evaluation. It also means that different process-centred approaches may be combined, such as Utilization-Focused Evaluation and Action Evaluation.

C. Coherence within the Overall Evaluation Plan

Focusing on the evaluation approach is somewhat akin to only considering the lungs when discussing health. The evaluation approach needs to be coherent with the evaluation objectives, the type of evaluation, the scope of the evaluation and the evaluators' role.³ The key evaluation questions raised within any given evaluation approach bear on the data collection and data analysis methods to be used. Evaluation approaches cannot simply be cut and pasted into different evaluation designs. As additional decisions are made about the design of the evaluation, it pays to check periodically to ensure the approach still corresponds to the other evaluation design considerations.

Selecting an evaluation approach does not automatically predetermine data collection⁴ or data analysis methods. Key informant interviews could be part of any approach and are not required

³ For more information on the components of an evaluation design see Church, Cheyanne, and Mark M. Rogers. *Designing for Results*. Search for Common Ground, Washington D.C. 2006. Chapters 8 & 9

⁴ For more information about data collection options in peacebuilding programs see: Carver, Christopher, and Rolf Sartorius. *Monitoring, Evaluation and Learning for Fragile States and Peacebuilding Programs*. Social Impact/ Office of Transition Initiatives (OTI), 2005. <http://www.manitouinc.com/PDFs/fragilestates.pdf>

by any single approach. Similarly, gap analysis could be used to analyze data regardless of the approach and is not a prerequisite to any particular evaluation approach.

II. Evaluation Approaches Useful in Peacebuilding

Evaluation approaches vary in their scope. Some, like Utilization-Focused Evaluation are so broad and comprehensive that it is difficult to summarize them in a few pages here. Others are much more circumspect and focused. All are still works in progress. Evaluation theorists and practitioners continue to publish their research and their insights stemming from practical experience.

Evaluation approaches are not entirely distinct from one another. Many engage participants in some step in the process, but this is different from Participatory Evaluation. Most approaches consider at some point the results stemming from the intervention. However, the types of results and the way in which results are evaluated differ from one approach to another. The overlaps and similarities can be confusing. It is important to understand the approach in its entirety. Avoid latching onto the first approach that meets your needs as several different approaches may address the same need.

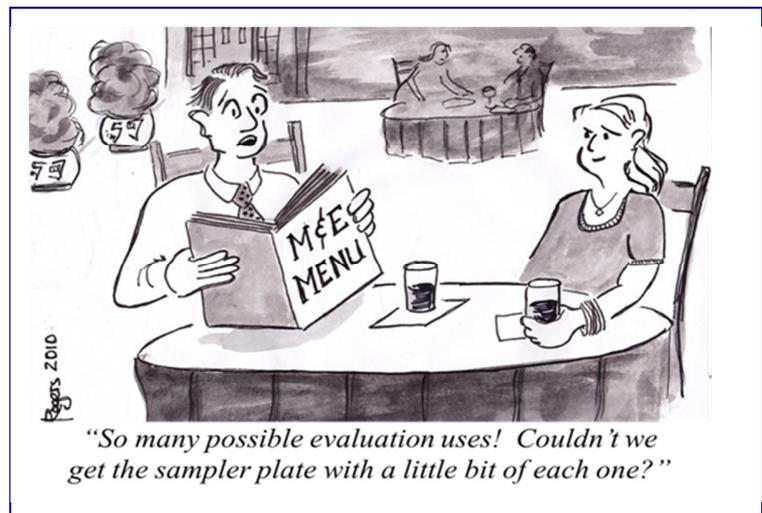
A. Process-centred Evaluation Approaches

The process-centred approaches tend to prioritize either the evaluation participants and decision makers or the eventual utilization of the evaluation over what is actually being evaluated.

Participatory Evaluation

Participatory Evaluation is an “educational process through which social groups produce action-oriented knowledge about their reality, clarify and articulate their norms and values, and reach consensus about further action.”⁵ This is very different from an external evaluation that involves programme participants in select steps in the evaluation process.

Participatory Evaluation encompasses several evaluation approaches, notably Action Evaluation and empowerment evaluation. All Participatory Evaluation approaches embrace a philosophy that evaluation data



⁵ Brunner, I., and A. Guzman, “Participatory Evaluation: A Tool to Assess Projects and Empower People.” In R. F. Conner and M. Hendricks (eds.), *International Innovations in Evaluation Methodology. New Directions for Evaluation*, No. 42. Jossey-Bass, San Francisco. 1989

are only valid when informed by practitioner perspectives.⁶ In the case of peacebuilding this may mean the protagonists, the practitioners of violence. Participatory Evaluation involves stakeholders in all aspects of the evaluation, including technical considerations.

The differences in the approaches to Participatory Evaluation involve the purpose of the evaluation, the diversity of the participants and the extent of their participation⁷, and the exercise of power within the evaluation process. For some approaches to Participatory Evaluation the purpose is to strengthen the programme and there is a relative balance in power between the evaluator and participants over the evaluation process. In more transformative approaches, the evaluation itself serves as a vehicle for empowerment and emancipation, putting ultimate decision-making in the hands of the programme participants.

Action Evaluation

Key Philosophies	Strengths	Weaknesses	Methods Implications
Iterative goal-setting process; Change is incremental Goal setting is a continuous process throughout implementation	Builds consensus on what to do and how Requires reflective practice Is adaptive	Requires regular periods of reflection and evaluation Time consuming Requires investment in consensus building	The “evaluator” plays the role of a learning facilitator Funder support is important from the beginning
	Potentially beneficial to use with Outcome Identification/Objective-based Study and Theory-based Evaluation		

Action Evaluation is designed for stakeholders to develop and periodically refine meaningful programme goals and corresponding evaluation criteria throughout the course of a peacebuilding initiative. Action Evaluation requires programme stakeholders to explicitly state and periodically revise their collective goals. Through a series of self-reflections, stakeholders determine what they wish to achieve and what success will look like. As they gain new experience and develop relationships and the context evolves, goals will need to be changed and better contextualized. This builds a sense of ownership and empowerment leading to greater commitment to the initiative. It also provides stakeholders with more nuanced criteria to consider in their ongoing evaluation of the initiative.

⁶ Levin, B. “Collaborative Research in and with Organizations.” *Qualitative Studies in Education*, Vol. 6, No. 4, 331–340. 1993

⁷ Daigneault, Pierre-Marc and Steve Jacob. “Toward Accurate Measurement of Participation: Rethinking the Conceptualization and Operationalization of Participatory Evaluation.” Published online *American Journal of Evaluation*, 10 July 2009

*Action Evaluation is a process of goal articulation and data gathering that is designed to systematize what is normally, though often haphazardly, done in design and implementation of most conflict resolution (and in many community and organizational development) interventions. This systematization of intervention process and content is intended to promote reflexive evaluation among all stakeholders as they move forward in their intervention.*⁸

Action Evaluation was very much in vogue within peacebuilding evaluation at the end of the 1990s and early 2000s. Action Evaluation predates more recent trends towards predetermined standards set externally and imposed upon programme participants through prescriptive requests for proposals and other funding mechanisms.

Action Evaluation shares a number of underlying philosophies with peacebuilding, including empowerment, self-determination, contextually adaptive action, collaboration and inclusion. Like peacebuilding, Action Evaluation values iterative processes as a means for dealing with complexity. It also encourages stakeholders to engage as reflective peacebuilding practitioners. Rothman maintains that, “Given that a great deal of conflict resolution is about raising awareness and enhancing interactive analysis--specifically about disputants' notions of and approaches to conflict--this process is very consistent with conflict resolution itself.”

Ross argues that Action Evaluation makes several crucial assumptions consistent with a great deal of social science theory and evidence⁹:

- The act of participation in the evaluation affects attitudes and behaviours
- Goal setting as an iterative process
- Change is an incremental process, hence goal setting is incremental
- The social construction of goals
- The interrelationship between theory and practice

These assumptions explain why Rothman advocates Action Evaluation as a way of building stakeholder awareness of and commitment to goals in peacebuilding initiatives.

⁸ Rothman, Jay. “Action Evaluation and Conflict Resolution Training: Theory, Method and Case Study.” *International Negotiation*, Vol. 2: 451-70. 1997

⁹ Ross, Marc Howard. “Action Evaluation in the Theory and Practice of Conflict Resolution,” *A Journal of The Network of Peace and Conflict Studies*, Vol. 8, No. 1, May 2001

Empowerment Evaluation

Key Philosophies	Strengths	Weaknesses	Methods Implications
Explicit interest in social justice	Mainstreams evaluation within programme cycle	May need to deal with differences between community and funder interests	Capacity needs have to be identified and addressed early
Stakeholder driven processes	Effective where staff are resistant to evaluation		Integrate with baseline and monitoring activities
Building evaluation capacity	Potentially beneficial to use with Outcome Identification/Objective-based Study		

Empowerment Evaluation, according to Wandersman, aims to increase the probability of achieving programme success by:

- 1.) providing programme stakeholders with tools for assessing the planning, implementation, and self-evaluation of their programme
- 2.) mainstreaming evaluation as part of the planning and management of the programme/organization.

Based on Wandersman, et al. (2005), the Centers for Disease Control and Prevention adapted the principles of Empowerment Evaluation to fit the violence prevention programmes it sponsors.¹⁰ Those principles are paraphrased below:

Community Ownership -- Empowerment Evaluation places the primary responsibility and ownership for building the organization's evaluation capacity and evaluating the organization's strategies with the organization, not the evaluator. The empowerment evaluator initially provides expertise, coaching, training, tools and technical assistance to the organization as it evaluates and builds its evaluation capacity. Eventually, organizational stakeholders have the capacity to conduct their own evaluations.

Inclusion - Empowerment Evaluation involves the representation and participation of key stakeholders.

Democratic Participation - Empowerment Evaluation is a highly collaborative process. Stakeholders are given the opportunity to voice questions, concerns and values throughout the evaluation process. Every stakeholder's voice is to be heard and valued equally.

¹⁰ Cox PJ, Keener D, Woodard T, Wandersman A. "Evaluation for Improvement: A Seven-Step Empowerment Evaluation Approach for Violence Prevention Organizations." Atlanta (GA): Centers for Disease Control and Prevention; 2009

Community Knowledge - Empowerment Evaluation values and promotes the knowledge present within violence prevention organizations and the communities within which they work. Organizational and community stakeholders, not external evaluators, are considered to be in the best position to understand the community's problems and to generate solutions to those problems.

Evidence-based Strategies - Empowerment Evaluation promotes the use of strategies with high-quality evidence of their strategies' effectiveness so that organizations can use their resources to select, implement and evaluate strategies that have a high likelihood of preventing violence. Evidence-based strategies are often complemented by community knowledge to ensure that a strategy is compatible with the community context.

Accountability - Empowerment Evaluation provides data that can be used to determine whether a strategy has achieved its goals. Negative results are not punished; rather, they are used to inform changes in a strategy or the selection of a new strategy for the purpose of producing better outcomes.

Improvement - Empowerment Evaluation helps organizations to improve their strategies so that they are more likely to achieve their stated goals and outcomes through activities such as needs assessments, assessments of the strategy's design, process evaluation and outcome evaluation.

Organizational Learning - Empowerment Evaluation fosters a culture of learning within organizations. Stakeholders come to view positive and negative evaluation results as valuable information that guides strategy improvement and to believe that every strategy can be improved.

Social justice - Empowerment Evaluation increases an organization's evaluation capacity to implement strategies that work to reduce disparities, discrimination, persecution, prejudice and intolerance.

Capacity building - Empowerment Evaluation builds individual and organizational evaluation capacity so that stakeholders are better able to conduct their own evaluations, understand results, and use them to continuously improve their strategies and their organization.

These principles are not exclusive to Empowerment Evaluation. Most evaluation approaches advocate evidence-based conclusions, accountability and learning. The social justice, democratic participation and capacity-building principles combined are largely what distinguish Empowerment Evaluation.

Utilization-Focused Evaluation

Key Philosophies	<i>Strengths</i>	<i>Weaknesses</i>	<i>Methods Implications</i>
Evaluation use is the driving concern	Helpful where there is resistance to evaluation	Time-consuming for participants	Participants need to be or become competent in the methods in which they participate
Up-front clarity on use increases acceptance	Well-known to development NGOs	Numerous steps within the overall process	
Intended use is continuously considered at every step	Well regarded in the evaluation community		Evaluators need considerable facilitation skills
Implicit philosophy of self-determination	Potentially beneficial to use with most other evaluation approaches including other process-centred approaches		

“Utilization-Focused Evaluation (U-FE) begins with the premise that evaluations should be judged by their utility and actual use; therefore, evaluators should facilitate the evaluation process and design any evaluation with careful consideration of how everything that is done, from beginning to end, will affect use.”¹¹ U-FE is a process. There is no specific content focus. There are no specific methods of data collection and analysis. U-FE does, however, adhere to certain principles:

- Situationally responsive
- Negotiated process between evaluators, stakeholders and other evaluation users
- Oriented toward facilitating decision-making about the issues being evaluated
- Involvement and engagement in the evaluation process encourage uptake of evaluation findings

Patton offers a checklist of a 17-step process. The checklist covers readiness or preparations, situation analysis, evaluation user identification and prioritization, evaluation questions and design, simulated use of fictitious findings, data collection, organization and reporting, follow-up with users and an evaluation of the evaluation. The steps are summarized below.¹²

1. *Assess and build programme and organizational readiness for UF-E* -- The evaluator considers the history of evaluation use, perceptions of evaluation, and incentives and barriers to evaluation use. This is also an opportunity to promote reality testing, establish

¹¹ Patton, Michael Quinn. *Essentials of Utilization-Focused Evaluation*, SAGE Publications Inc., Thousand Oaks, CA. 2012

¹² Patton, Michael Quinn. *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*, The Guilford Press, New York. 2010

project-specific norms and standards, introduce professional evaluation standards and create a positive image for evaluation.

2. *Assess and enhance evaluator readiness and competence to undertake a utilization-focused evaluation* – Those commissioning the evaluation and the evaluator need to consider the evaluators' competence, commitment to evaluation use and their credibility. Key competencies reach beyond technical and professional norms, standards and principles. They also involve being able to understand and attend to “the contextual and political issues of an evaluation, including determining evaluability, addressing conflicts and attending to issues of evaluation use.” Evaluators need interpersonal competence and project management skills.
3. *Identify, organize and engage intended users* – Stakeholder analysis is frequently a key component of conflict analysis and is very similar to this step in U-FE. The idea is to engage the “strategic who” - those with high levels of interest and influence and those who represent key constituencies. Success depends on facilitating high-quality interactions between a manageable number of primary intended users and nurturing their interest in and capacity to engage in evaluation.
4. *Situation analysis conducted jointly with primary intended users* – Here Patton refers to the evaluation situation, particularly as it relates to evaluation use. Peacebuilding programmes have the additional requirement of conducting a conflict analysis and ensuring that the evaluation is conflict sensitive.
5. *Identify and prioritize interested users by determining primary purposes* - This is an integrative task involving coherence between the evaluation purposes (see six options mentioned earlier on page 3) and intended evaluation users with what is at stake. For example, an evaluation to be used by funders to make an overall summative judgement regarding the future of the programme would be very high stakes. By contrast, a formative evaluation to be used by programme staff and participants to modify and improve specific implementation practices might have moderate stakes.
6. *Consider and build in process uses if and as appropriate* – “Process use occurs when those involved in the evaluation learn from the evaluation process itself or make changes based on the evaluation process rather than just the evaluation's findings.” Patton has identified six varieties of process use, notably infusing evaluative thinking into the organizational culture and incorporating evaluative questioning into decision making.
7. *Focus priority evaluation questions* – “Focusing an evaluation is an interactive process between evaluators and the primary intended users of evaluation.” Since it is not possible to evaluate everything, strategic choices need to be made. Patton lists over 50 ways of focusing or types of evaluation ranging from Appreciative Inquiry to Theory of Change.
8. *Check that the fundamental areas for evaluation inquiry are being adequately addressed* – The evaluation fundamentals centre on implementation, outcomes and attribution. “The three classic questions are:

- The implementation questions: What happened in the programme?
- The outcomes question: What were the results?
- The attribution questions: Can what resulted be attributed to what was implemented?"

Regardless of the purpose and focus of an evaluation, some consideration also needs to be given to these evaluation fundamentals.

9. *Determine if the intervention model or theory of change is being evaluated* – This involves articulating the intervention model and the underlying theories on which it rests. The programme logic is illustrated through a hierarchy or chain of results. (For more information see Theory-based Evaluation below.)
10. *Negotiate appropriate methods to generate credible findings that support intended use by intended users* – The choice of methods is dependent on the purpose of the evaluation, the intended use of the findings, the resources for the evaluation and the trade-off negotiated during the evaluation design. U-FE has 10 methods/principles dealing with validity, data quality, utility, credibility and the dynamics of design decisions and measurement.
11. *Make sure the intended users understand the potential methods controversies and their implication* – The evaluator needs to provide the intended evaluation users with information about the strengths and weaknesses and any surrounding controversies relating to the methods options. This includes the benefits and costs of using mixed, both qualitative and quantitative methods.
12. *Simulate the use of findings* – Patton equates this with a dress rehearsal. The evaluator prepares a fictitious data set and findings, and the intended evaluation users simulate trying to make decisions or using the evaluation findings. Findings that are not useful can be modified or replaced with better options before the expense of actual data collection.
13. *Gather data with ongoing attention to use* – Primary intended users need to stay informed and involved, including during the data collection.
14. *Organize and present the data for interpretation and use by the intended users* --
“Facilitating data interpretation among the primary intended users increases their understanding of the findings, their sense of ownership of the evaluation, and their commitment to use the findings.”
15. *Prepare an evaluation report to facilitate use and disseminate significant findings to expand influence* – Dissemination may also foster evaluation use and expand the sphere of influence.

16. *Follow up with primary intended users to facilitate and enhance use* – The intent here is to help evaluation users understand evaluation as an ongoing process rather than a one-time event.
17. *Meta-evaluation of use: Be accountable, learn and improve* – Meta-evaluation refers to the evaluation of an evaluation. In the case of U-FE, meta-evaluation also considers the use and utility of the evaluation. Meta-evaluation is most effective when undertaken at the same time as the evaluation, rather than after the programme and the evaluation are over.

The U-FE process is not as linear as it appears. Several of the early steps may be concurrent. Patton struggled with the sequencing of steps 7 and 8. A case could be made for either preceding the other.

Goal-free Evaluation

Key Philosophies	<i>Strengths</i>	<i>Weaknesses</i>	<i>Methods Implications</i>
Evaluator objectivity and independence Inductive thinking based on empirical evidence	Helpful in clarifying programme goals Holistic Detects positive and negative intended and unintended outcomes	May substitute the evaluators' bias in place of the programme's bias	Requires two evaluators, at least one of whom must be external
Potentially beneficial to use with Theory-based Evaluation, Outcome Identification and Objectives-based Studies			

In this approach, an external evaluator attempts to identify a programme's results or outcomes with no prior knowledge of the programme design. The idea is to judge the extent to which the actual outcomes meet participants' needs. By limiting exposure to the goals and the discussions with staff about the goals, the evaluation is not biased by the orientation of the programme. This helps prevent missing unanticipated outcomes and maintain evaluator objectivity and independence.

The evaluation begins with an open-ended inquiry into the changes in the conflict or the context over the time period of interest. Other approaches to programme evaluation typically seek to know what changes have occurred as a result of the intervention. Goal-Free Evaluation asks first about changes and secondly about what people think about how and why those changes took place. In Goal-Free Evaluation of a peacebuilding programme, the evaluation begins with a conflict analysis – what is the conflict and how has it changed? The intervention may figure among multiple answers, as may other interventions, dynamics or environmental considerations.

When used as a precursor to theory-based evaluation, Goal-Free Evaluation can strengthen claims of the intervention's contribution to a given change or effect. It validates the outcomes and effects and can help identify other theories that explain the interaction between the effect and the intervention.

Goal-Free Evaluation is rare in development circles and peacebuilding. It requires at least one external evaluator, which can drive up the cost. Controlling for the programme's potential to bias does little to keep in check other sources of bias such as those the evaluator brings to the process. Programme managers sometimes perceive Goal-Free Evaluation to be somewhat risky.

B. Change-centred Approaches to Evaluation

Other evaluation approaches focus on the changes a programme seeks to bring about. These may be at the level of implementation, programme outcomes or changes in the conflict attributable to

the intervention, also known as impact. Given the large number of NGO M&E guides dealing with implementation monitoring and evaluation, it is not covered here. The key question about implementation is: “Did we do what we said we would do?”

Outcome Identification/Objectives-based Studies

Key Philosophies	Strengths	Weaknesses	Methods Implications
Explores whether or not select objectives were achieved	Well known	Not useful in judging programme merit	Need to compensate for narrow focus
	Facilitates focus on a specific component of a programme	Blind to other dynamics in play (i.e. unintended consequences)	
	Manageable		
Potentially beneficial to use with Theory-based Evaluation, Impact Evaluation and Implementation Evaluation			

This approach focuses on objective-level changes, or outcomes. Stufflebeam refers to these as objective-based studies characterized by a question orientation as opposed to a methods orientation. “The questions-oriented approaches usually begin with a set of narrowly defined questions. These might be derived from a programme’s behavioural/operational objectives, a funding agency’s pointed accountability requirements, or an expert’s preferred set of evaluative criteria.”

Evaluations looking exclusively at outcomes are better suited to address monitoring, programme development and learning needs or purposes rather than determining a programme’s merit or worth. According to Stufflebeam, “The objectives-based approach is especially applicable in assessing tightly focused projects that have clear, supportable objectives. Even then, such studies can be strengthened by judging project objectives against the intended beneficiaries’ assessed needs, searching for side effects, and studying the process as well as the outcomes.”

Outcome Identification and Objective-based Studies tend to focus on questions¹³ such as:

- What intended outcomes occurred?
- Did the outcomes align with expectations?
- What part of the project was the most important in catalyzing change?
- What was the process (environment, community, personal) that catalyzed the shift?
- What prior changes were required before this outcome could occur?

“The objectives-based study has been the most prevalent approach in programme evaluation. It has commonsense appeal; programme administrators have had a great amount of experience with it; and it makes use of technologies of behavioural objectives,

¹³ Church, Cheyanne, and Mark M. Rogers. *Designing for Results*. Search for Common Ground, Washington D.C. 2006

both norm-referenced and criterion-referenced testing, and performance assessments. Common criticisms are that such studies lead to terminal information that is neither timely nor pertinent to improving a program's process; that the information often is far too narrow to constitute a sufficient basis for judging the object's merit and worth; that the studies do not uncover positive and negative side effects; and that they may credit unworthy objectives." (Stufflebeam)

Two of the better-known processes within this approach include Most Significant Change¹⁴ and Outcome Mapping.¹⁵ The Most Significant Change (MSC) technique is a form of participatory monitoring and evaluation that involves the collection of stories at the field level that are systematically analyzed to identify project outcomes and changes in the conflict. MSC provides a method for capturing and analyzing stories and exploring values behind the preferences for certain changes. MSC may provide programme stakeholders and participants with a better understanding of what is and is not being achieved. Because of its open-ended questions, data can be collected about multiple dynamics or the overall project as a whole, rather than just the intended outcomes.

Outcome Mapping (OM) is a method used to create planning, monitoring and evaluation mechanisms that enable programmes to document, learn from and report on their results. It is designed to promote understanding a programme's achievements, while recognizing the contributions of other actors that are essential to achieving macro-level changes. In OM, the shift is "away from assessing the development impact of a programme (defined as changes in state — for example, policy relevance, poverty alleviation or reduced conflict) and toward changes in the behaviours, relationships, actions or activities of the people, groups and organizations with whom a development programme works directly."

Outcome Mapping is divided into three stages:

1. *Intentional Design*-- designates the intended macro-level changes and corresponding strategies.
2. *Outcome and Performance Monitoring*-- sets a self-assessment framework and data collection tools for the ongoing monitoring of the programme's actions and progress towards results.
3. *Evaluation Planning*-- sets the evaluation priorities and develops an evaluation plan.

Outcome Mapping recognizes that multiple, nonlinear events lead to change. OM looks at the logical links between interventions and behavioural change. Outcome Mapping assumes only that a contribution has been made, rather than assuming or attempting to claim attribution.

¹⁴ Dart, Jess and Rick Davies. *The 'Most Significant Change' (MSC) Technique: A Guide to Its Use*. Cooperative for Assistance and Relief Everywhere (CARE) Inc., et al., 2005.

¹⁵ International Development Research Centre. http://web.idrc.ca/en/ev-26586-201-1-DO_TOPIC.html

Theory-Based Evaluation¹⁶

Key Philosophies	Strengths	Weaknesses	Methods Implications
Looks at why and how lower-level results leverage higher-level changes	Makes the programme logic explicit	Most compelling when implementation is flawless and outcomes have been achieved	Dependent on good evidence of the results within the theory(ies) being evaluated
	Can be used to assess programme logic prior to implementation		
Compatible with Impact Evaluation, Outcome Identification/Objectives-based Study and Implementation Evaluation			

Theory-Based Evaluation (TBE) explores the transformation of inputs and outputs into results, focusing on how and why change occurs. Theory-Based Evaluation seeks to understand why programmes do and do not work so that programme developers can create effective, sensible and coherent programmes. According to Weiss, Theory-Based Evaluation can help distinguish poor theory from poor implementation.

Why use Theories of Change?¹⁷

- Establish common principles and vocabulary
- Make implicit assumptions explicit
- Identify resources and check them for adequacy
- Design more realistic plans of action
- Clarify lines of responsibility
- Create more meaningful evaluation and “maintain healthy scepticism”

Stufflebeam (2001) identified numerous questions addressed by TBE. These include:

- Is the programme grounded in appropriate, well-articulated and validated theory?
- Does the employed theory reflect recent research?
- Are the programme’s targeted stakeholders, design, operation and intended outcomes consistent with the guiding theory or theories?
- Are programme inputs and operations producing outcomes in the ways the theories predict?
- What changes in the programme’s design or implementation might produce better outcomes?
- What elements of the programme are essential for successful replication?
- Overall, was the programme theoretically sound? Did it operate in accordance with an appropriate theory?
- Is the programme worthy of continuation and/or dissemination?

¹⁶ This section on theory-based evaluation is an excerpt from Rogers, Mark M. "Theory-Based Evaluation." Unpublished paper submitted to the U.S. Agency for International Development (USAID/DCHA/CMM), Washington, D.C. 2010

¹⁷ “Mapping Change: Using a Theory of Change to Guide Planning and Evaluation.” GrantCraft, 2006 (www.grantcraft.org/index.cfm?fuseaction=Page.ViewPage&pageId=1542)

Although TBE is most frequently used during and after implementation when empirical evidence is available, it may also be used to explore the programme logic at the design state. This is done by assessing the plausibility of the theories, their centrality to the problem or need, and any ethical issues implicit within the theories.

To evaluate theories in practice by a programme already under way, they need to be at the point where key activities have taken place and outcomes of interest have been manifest.¹⁸ Internal validity includes clarity, connections to the problem and plausibility. External validity has to do with the relevance to the context, being consistent with known theory and any ethical bias.¹⁹

There are a number of instances, particularly relevant to peacebuilding, where Theory-Based Evaluative evaluation. These include situations



“Dear Mr. Gandhi, We regret we cannot fund your proposal because the link between spinning cloth and the fall of the British Empire was not clear to us.”

Written by M. Rogers and Illustrated by Ariv R. Faizal, Wahyu S. Ary W.S.
Creative Team Search for Common Ground in Indonesia

- The context has changed since the programme began
- Programmes are conceived in haste under challenging circumstances
- There have been rapid dramatic shifts in programming
- Assumptions within innovative or pilot programmes need to be validated or revalidated

As Weiss points out, Theory-Based Evaluation is not appropriate in every circumstance. For example, a security

sector reform programme does not need to use TBE to revalidate the effectiveness of reducing the soldier/commander ratio in improving command control. By the same token, in dialogue programmes, there is no need for TBE to reconfirm the value of one group validating the interests of the other group in improving communications between the groups.

Theory-Based Evaluation need not test every programme theory in a given programme. Many peacebuilding programmes address complex issues and contain multiple programme theories of change and even multiple stands of lower-level theories within any given programme theory. Looking at them all in the course of a single evaluation is rarely realistic or desirable. Carol H. Weiss²⁰ identified four criteria to consider in determining priorities or the focus of Theory-Based Evaluation:

- *Plausibility* – Can the programme actually achieve the results the theory assumes?

¹⁸ Bickman, Leonard. “Summing Up Program Theory.” *New Directions in Evaluation*. 87:103-112. 2000

¹⁹ For more information on evaluating program theories of change, see Funnel, Sue and Patricia J. Rogers, *Purposeful Program Theory*, Jossey-Bass, San Francisco, 2011

²⁰ Weiss, Carol H. “Which Links in Which Programs Shall We Evaluate?” *New Directions in Evaluation*. 87:35-46. 2000

- *Centrality to the programme* – Some theories are so essential that, no matter what else happens, the programme’s success hinges on the viability of this particular theory.
- *Uncertainty* – This involves the lack of knowledge in the programme field, or the weaker elements of the programme logic.
- *Beliefs of the programme stakeholders* – What pathways do stakeholders assume lead to good outcomes?²¹

Where theories of change are implicit or unarticulated, TBE may benefit from the participation of an external evaluator. Because TBE hinges on the clarity and strength of the theories of change, it is best served by evaluators with knowledge of the subject matter and TBE. TBE is resource intensive and it is most convincing when used in conjunction with other evaluation approaches such as Outcome Identification/measurement and Implementation Evaluation. [

Other evaluation approaches, such as Impact Evaluation, may be more appropriate for determining whether or not changes in the conflict can be attributed to an intervention. TBE’s usefulness in determining causality is limited to specific situations. However, it is difficult to do well (Stufflebeam 2001). Cook prefers “...to see theory-based methods used within an experimental framework and not as an alternative to it.” (2010) Others, including Weiss and Patricia Rogers, believe that Theory-Based Evaluation can be valuable independent of experimental designs. Given the nascent nature of the field of peacebuilding and the challenges of using experimental designs in conflict environments, Theory-Based Evaluation, even with its challenges and limitations, has much to offer.

²¹ In conflictive environments, some evaluators distinguish different theories according to the points of view of different stakeholder groups, rather than trying to develop or assume a consensus around a common theory of change. For more information see Morten, Hansen and Evert Vedung. *Theory-Based Stakeholder Evaluation*, American Journal of Evaluation Vol. 31, No. 3. SAGE Publications Inc., Thousand Oaks, CA. 2010

Impact Evaluation

Key Philosophies	<i>Strengths</i>	<i>Weaknesses</i>	<i>Methods Implications</i>
Seeks to attribute changes in the conflict to programme interventions	Can be convincing and useful in policy development and learning	Not appropriate for many conflicts	Fairly sophisticated experimental, quasi-experimental and in select cases non-experimental (naturalistic) methods
Causation is demonstrated when all other intermediating variables are controlled	Potentially beneficial to use with Theory-based Evaluation		

Impact Evaluation attempts to determine the changes in the conflict or context that are attributable to a given intervention. Establishing causality is at the heart of Impact Evaluation. The attribution question is what separates Impact Evaluation from the other approaches.

Impact is relative, which explains why it is used in so many different ways. For many programmes, impact is reserved for goal-level changes, such as effects on *Peace Writ Large*. Programmes with less ambitious aspirations might have local changes as goals they consider as impacts.

Timing an Impact Evaluation largely depends on when one expects the changes to be manifest. For some, impact is reserved for longer-term changes, such as post-programme changes and, hence, should be evaluated only after the programme is over. In other cases, there may be short-term impacts of consequence that merit evaluation while the programme is still under way. Evaluability assessments are helpful in preparing for and deciding whether or not Impact Evaluation is appropriate. Evaluability assessments generally look at the programme design and scope, contextual issues, timing, documentation and evidence on hand and the potential learning return on the investment in Impact Evaluation. Impact Evaluation is most appropriate when:

- the activities have been well implemented
- the output and outcomes are well-evidenced
- there is clear programme theory that explains how the interventions influenced the changes in the conflict
- stakeholders need to determine if the changes in the conflict are attributable to the programme.

Patricia Rogers and the Network of Networks of Impact Evaluation (NONIE) Subgroup 2 (2008) maintain that Impact Evaluations include the following tasks:

- *Comprehensive identification of important impacts* -- changes in the context and/or conflict, whether negative, positive, intended and unintended, and their durability over time. In practice, this amounts to updating or completing a conflict analysis. Impact Evaluation begins by considering all changes in the conflict, not just those specified

within the intervention.

- *Systematic and defensible data collection and analysis of evidence of these impacts* – In the past there was much debate about the best methods of data collection and analysis for use in Impact Evaluation. Today the general consensus is that mixed, qualitative and quantitative methods are needed in Impact Evaluation.
- *Sound inferences about the contribution of the intervention to achieving these impacts* -- Inference is used to explain the causal links between an intervention and an effect. Alternative explanations need to be considered and either eliminated or controlled for where possible. The use of counterfactual, or control groups, can be effective where the programme can be manipulated in order to control specific variables. Quasi-experimental methods²² try to approximate or mimic either randomness or representation. There are also qualitative methods for inferring causal relationships between an intervention and an effect.²³
- *Effective management of the evaluation*, including transparent reporting of methodology threats to validity, and where appropriate, formal evaluation of the evaluation (also known as meta-evaluation).

Many peacebuilding programmes request Impact Evaluation when they mean Outcome Identification or Objectives-based Studies. Peacebuilding programmes charged with determining attribution often recognize the related challenges and constraints, and prefer to focus on contribution rather than attribution. Evaluation guru Michael Quinn Patton maintains (2008) that contribution analysis is particularly appropriate in situations where multiple initiatives are working toward the same end, where impacts occur over long periods of time, and where “simple notions of linear attribution are neither meaningful nor accurate.” Unfortunately, contribution analysis is a recent (2001) and still very underdeveloped methodology. It hinges on providing a preponderance of evidence that allows one to draw a reasonable conclusion. There is no consensus within the peacebuilding field on what types of change are essential much less what evidence is needed to determine whether or not a change was influenced by a specific intervention among other interventions and other related factors.

Impact Evaluation requires a substantial investment of resources, qualified evaluators and situationally appropriate mixed methods. The opinion of a learned and connected resource person does not qualify as Impact Evaluation. Nor is Impact Evaluation a secondary, add-on to

²² Oxfam is looking at quasi-experimental designs. See Hughes, Karl and Claire Hutchings, “Can we obtain the required rigour without the randomisation? Oxfam GB’s non-experimental Global Performance Framework.” International Initiative for Impact Evaluation Working Paper #13, New Delhi, August 2001

²³ Patton (2010) offers empirical data with source triangulation as a viable alternative to experimental and quasi-experimental options. Scriven offers *modus operandi* and *causal list* methods to argue for or against attribution. Miles and Huberman (1994) provide a more nuanced understanding and offer an expanded taxonomy of nine different types of causal relationships between variables. Qualitative methods rely on multiple lines and levels of evidence.

another evaluation priority. It is challenging enough in its own right, and more so in peacebuilding. That said, Impact Evaluation is most convincing when supported with Theory-Based Evaluation.

Programme managers considering Impact Evaluation should specifically ask the evaluators to describe experimental, quasi-experimental and non-experimental (naturalistic) options for evaluating the intervention of interest.

Developmental Evaluation

Key Philosophies	<i>Strengths</i>	<i>Weaknesses</i>	<i>Methods Implications</i>
Programmes exploring ways to address systems in complex dynamic situations benefit from rapid, early evaluation feedback to improve emergent designs	Flexible and adaptive Useful where there are no known solutions	Easily confused with formative or pre-formative evaluation Requires close collaboration between developers & evaluators	No counterfactuals Looks for unanticipated results Reliance on inference
	Potentially beneficial to use with Outcome Identification/Objectives-based Study and Implementation Evaluation		

Developmental Evaluation is perceived as one of the new approaches, but has actually been in practice for years. Developmental Evaluation refers to the application of evaluative thinking and evaluation methods and processes to cutting edge, iterative and emergent social innovation. Developmental Evaluation is not a distinct evaluation process. Nor does it focus on specific content or issues. It is more about when the evaluation is conducted (as early as possible) and its purpose (feedback into the creative/conceptualization process) and the context in which the work is done (complex dynamic systems). It is largely for innovative and experimental initiatives before there is a programme model. It also serves in dynamic situations where continuous and on-going development and adaptation are the goal rather than arriving at a final model, or taking a model to scale.

The purpose of Developmental Evaluation is to provide timely feedback and generate learning and support action during the process of programme development. Developmental Evaluation focuses on systems changes along a continuum from small local systems to disruptive social innovations aimed at major, cross-scale impact on big problems. Developmental Evaluation is most appropriate for complex and dynamic environments where there are no known solutions to priority problems, where there is no certain way forward and multiple pathways are possible, and where there is a need for exploration and social experimentation.

In this sense, Developmental Evaluation is well suited for rapidly adaptive peacebuilding initiatives. However, few donors fund such undefined and open-ended initiatives, which by definition lack preconceived deliverables and pose accountability challenges.

Developmental Evaluation is characterized by the following:²⁴

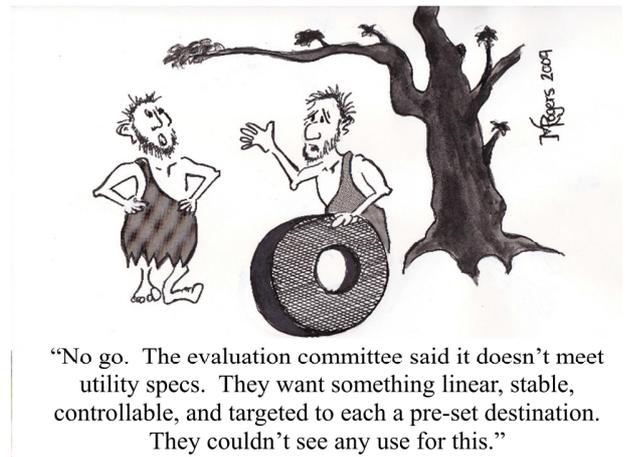
- The interpretive and reasoning processes include inference, pragmatism and contribution analysis.
- DE requires methodological flexibility and adaptability. It requires systems thinking, creativity, critical thinking and a high tolerance for ambiguity.
- Counterfactual formulations are meaningless because there are far too many variables and possibilities emerging and interacting dynamically to conceptualize simple counterfactuals.
- Serious attention is given to the unanticipated and the emergent as a fundamental evaluation function.

Developmental Evaluation brings the evaluator into the process of programme development, testing and refinement. While many programmes may have originated as adaptive, responsive designs, few funders underwrite the kind of learning laboratory where Developmental Evaluation has the most to offer.

Thinking about Evaluation Approaches

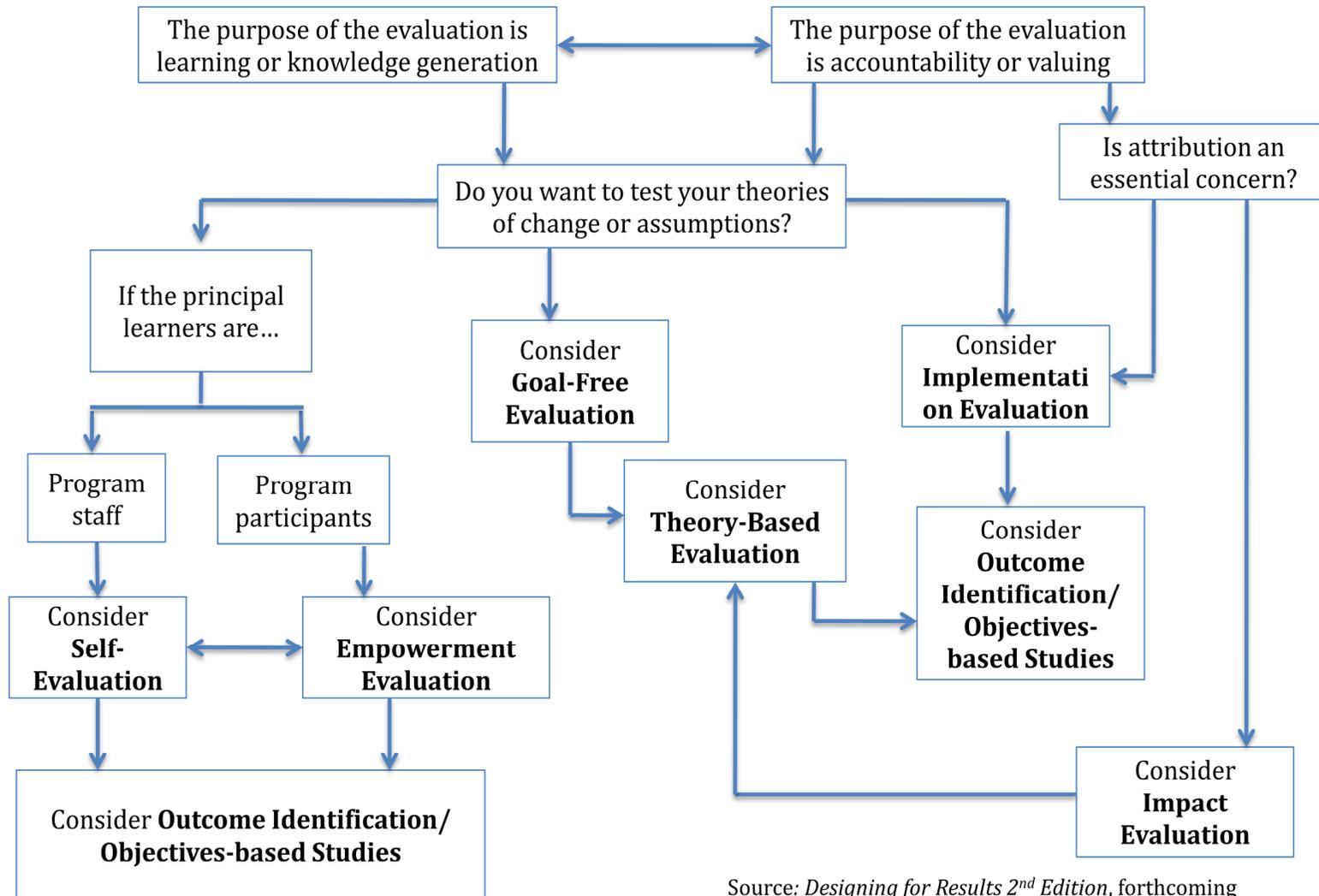
The flow chart on the following page attempts to illustrate how some of the different approaches might be used in combination. Regardless of the purpose of the evaluation, most evaluations still have to deal with the essentials, particularly questions about results or outcomes. Reflective practices, such as Action Evaluation, still need to pay heed to and consider the outcomes or results of their work.

Evaluator and theorist Michael Scriven is fond of saying, “evaluation is never complete.” In part this is due to the practice of bounding or limiting the work into manageable parts. The challenge is making sure the parts that support and depend on each other are not separated, or at least not without due consideration.



²⁴ Patton, Michael Quinn. *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*, The Guilford Press, New York. 2010

Thinking about Evaluation Approaches



Source: *Designing for Results 2nd Edition*, forthcoming

III. Cost Considerations

Cost is not a useful criterion for choosing one approach over another. Cost is often very important in selecting the evaluation methods within a given approach. There are too many variables involved in designing evaluations to be able to provide actual costs for comparisons between different evaluation approaches.

There are a range of variables affecting cost including:

- Scope of the intervention
- Depth, thoroughness and date of usable conflict analysis
- Intended generalizability of the evaluation findings
- Methodological rigour in the data collection and data analysis
- Number and type of evaluation objectives (i.e., OECD-DAC criteria)
- Utility and reliability of existing data
- Complexity of the conflict and the context
- Complexity and number of the interventions
- Role of external evaluators (if any)
- The importance of attribution

Generalizability refers to inferences made about larger groups or populations after evaluating only a portion of the group or population. Random probability sampling ensures the representativeness of a smaller group or subgroup needed to draw conclusions about the larger population. For quantitative evaluation methods, this often is referred to as statistical inference.

Purposeful sampling in qualitative inquiry permits in-depth insights that are difficult to achieve through random probability sampling. There are over a dozen other strategies for selecting samples for evaluation, such as convenience, critical case, deviant case, maximum variation and opportunistic sampling. Many breakthroughs in social science came from studying a few individuals, hence the value of case studies.

Evaluations seeking to make generalizations based on random sampling may be more expensive because of the volume of data to collect and the logistics needed to reach numerous and sometimes remote areas selected by random.

Rigour refers to the validity and reliability of information obtained through a given methodology. Generally speaking, validity refers to accuracy and how well the measures used reflect the concept or change they are intended to measure. For example, asking whether or not increases in civilian mobility and movement within a conflicted area really measures changes in trust is a validity question. “Reliability is defined as the consistency of information and information collection procedures.”²⁵

²⁵ Yarbrough, Donald B., et. al. *The Program Evaluation Standards*, SAGE Publications Inc., Thousand Oaks, CA, 2011

Methodological rigour can be thought of as a continuum ranging from pragmatic to scientific. There are choices along the continuum. Research leans toward the scientific end. Evaluation tends to be more pragmatic. Patton explains, “Judgements about validity and reliability, for example, are necessary and appropriately relative rather than absolute in that the rigor and quality of an evaluation’s design and measurement depend on the purpose and intended use of the evaluation.”²⁶ The options invariably have cost considerations. Generally speaking greater rigour implies greater cost.

Complexity of the conflict covers a range of variables such as the number and type of stakeholders, lethality and duration of the conflict, root causes, access to stakeholders, role of the Diaspora, safety of the evaluation participants, evaluators and staff, etc. One would not expect an evaluation of an NGO intervention in the Falkland conflict, for example, to compare with one in the Israeli-Palestinian conflict.

Actual costs will vary, sometimes even within one country. In the following table, different cost options are compared in relative terms rather than in monetary terms. Few evaluations fall strictly into a single column. More commonly there is a mix, where one element has been reduced in order to have resources available for more detailed or rigorous work on another concern. For example, a programme with multiple partners may reduce the number of partners engaged in the evaluation in order to be able to cover a greater number of the OECD-DAC criteria. Or a programme may choose to trade off generalizability (being able to extrapolate conclusions to other places) by using purposeful sampling rather than random sampling in order to be able to apply mixed methods in fewer sites.

As implied by the table, there are ways to keep evaluation costs in check:

- Maintain an accurate, up-to-date conflict analysis
- Ensure high quality and relevant monitoring information
- Bound or focus the evaluation on the most important use/learning
- Develop and integrate the key evaluation questions as part of the programme design
- Engage in joint evaluation with similar or related programmes

For detailed information on how to develop a budget for an evaluation see *Designing for Results*.

²⁶ Patton, Michael Quinn, *Utilization-Focused Evaluation: The New Century Text*, 4th edition. SAGE Publications Inc., Thousand Oaks, CA, 2008

Cost Variables	Shoestring Budget	Moderately Resourced	Very Well Resourced
Conflict Analysis	High quality recent conflict analysis on hand	Easily updated quality conflict analysis on hand	No usable conflict analyses on hand
Scope	Few sites in limited area	Large numbers of sites in a limited area	Large numbers of sites in multiple areas
Partnership/Interventions	0-2 partnerships/interventions	2-4 partners/interventions	5 or more partners/interventions
Generalizability	Limited to local context	Generalizable to neighbouring sites	Generalizable to national level or beyond
Rigour/methods	Less rigorous methods	Rigorous but not optimal methods	Optimal rigorous methods
Number and type of evaluation objectives/OECD-DAC criteria	1-2 criteria	3-4 criteria	5-7 criteria
Utility and reliability of existing data	Manageable amounts of highly useful and reliable data on hand	Limited useful and credible data on hand	Large volumes of unrelated data on hand or no data on hand
Accessibility to the evaluand (that which is being evaluated)	Easily and quickly accessed on multiple occasions using various means	Access depends on hard-to-get authorisations and limited availability	Limited, difficult access, or remote/virtual access
Extent to which the evaluation serves to build evaluation capacity of the evaluand	Draws upon existing high levels of evaluation capacity	Limited windows of learning by doing (data collection and analysis)	Linked to formal training in evaluation and practice
Complexity of the context	Primarily complicated	Mixture of complicated and complex	Primarily complex and dynamical
Complexity and number of the interventions	Few, simple and complicated interventions	Few multifaceted interventions	Multiple complex interventions
Role of external evaluator	Advisor	Operative	Operative
Importance of Attribution	Limited consideration of contribution	Thorough consideration of contribution	Seeks to establish attribution

IV. Conclusions

Choosing an evaluation approach or approaches is inherently a strategic process. Rarely are these choices made well in isolation. Technical assistance is often warranted. Where M&E resource people are on staff or accessible in regional or country offices, it makes sense to invite them into the process at the conception stage. External evaluators submitting tenders should justify the approach(es) they propose. For ambitious and expensive evaluations, country programmes may wish to retain an external evaluator to coach and support programme managers in their discussions and negotiations with the evaluation teams who will be implementing the evaluation.

V. Resources

Bamberger, Michael, Mary Church, Lucia Fort, and Jim Rugh. "Shoestring Evaluation: Designing Impact Evaluations under Budget, Time and Data Constraints." *American Journal of Evaluation* Vol. 25, No. 1: 5–37. 2004

Bickman, Leonard. "Summing Up Program Theory." *New Directions for Evaluation*. Issue 87:103-112. 2000

Brunner, I., and Guzman, A. "Participatory Evaluation: A Tool to Assess Projects and Empower People." In R. F. Conner and M. Hendricks (eds.), *International Innovations in Evaluation Methodology*. *New Directions for Evaluation*, No. 42. Jossey-Bass, San Francisco. 1989

Carver, Christopher, and Rolf Sartorius. *Monitoring, Evaluation and Learning for Fragile States and Peacebuilding Programs*. Social Impact/ Office of Transition Initiatives (OTI), 2005

Chambers, Robert, *et al.* "Designing Impact Evaluations: Different Perspectives," International Initiative for Impact Evaluation Working Paper #4, New Delhi, July 2009

Chen, Huey-Tsyh. *Theory-Driven Evaluations*. SAGE Publications Inc., Thousand Oaks, CA. 1990

Church, Cheyanne, and Mark M. Rogers. *Designing for Results*. Search for Common Ground, Washington D.C. 2006

Connell, J. P, A. C. Kubisch, L. B. Schorr, and C. H. Weiss. *New Approaches to Evaluating Community Initiatives: Concepts, Methods and Contexts*. Washington D.C.: The Aspen Foundation, 1994

Cook, Thomas. 2010. "The False Choice Between Theory-Based Evaluation and Experimentation." *New Directions for Evaluation*. No. 87:27-34. Fall 2000

Cousins, J. B., and Whitmore, E. "Framing Participatory Evaluation." In E. Whitmore (ed.), *Understanding and Practicing Participatory Evaluation*. *New Directions for Evaluation*, No. 80: 5-23.1998

Cox PJ, Keener D, Woodard T, Wandersman A. "Evaluation for Improvement: A Seven-Step Empowerment Evaluation Approach for Violence Prevention Organizations." Centers for Disease Control and Prevention, Atlanta, GA. 2009

Daigneault, Pierre-Marc and Steve Jacob. "Toward Accurate Measurement of Participation: Rethinking the Conceptualization and Operationalization of Participatory Evaluation." Published online *American Journal of Evaluation*, 10 July 2009

Dart, Jess and Rick Davies. *The 'Most Significant Change' (MSC) Technique: A Guide to Its Use*. Cooperative for Assistance and Relief Everywhere (CARE) Inc., et al, 2005

- Davidson, Jane, E. "Ascertaining Causality in Theory-Based Evaluation," *New Directions for Evaluation*. No. 87:17-26. 2010
- Gertler, Paul J., et. al., *Impact Evaluation in Practice*, The International Bank for Reconstruction and Development / The World Bank, Washington, D.C. 2011 <http://www.worldbank.org/pdt>
- Hansen, Morten and Evert Vedung. "Theory-Based Stakeholder Evaluation." *American Journal of Evaluation* Vol. 31, No. 3:295-313. 2010
- Hughes, Karl and Claire Hutchings. "Can we obtain the required rigour without the randomisation. Oxfam GB's non-experimental Global Performance Framework." International Initiative for Impact Evaluation Working Paper #13, New Delhi, 2001
- International Initiative for Impact Evaluation (3ie), "Principles for Impact Evaluation," undated, <http://www.3ieimpact.org/strategy/pdfs/principles%20for%20impact%20evaluation.pdf>
- Leeuw, Frans and Jos Vaeseen. *Impact Evaluations and Development: NONIE Guidance on Impact Evaluation*. Washington D.C.: NONIE—The Network of Networks on Impact Evaluation, 2009
- Levin, B. "Collaborative Research in and with Organizations." *Qualitative Studies in Education*, Vol. 6, No. 4, 331–340. 1993
- Marquart, Jules. "A Pattern-Matching Approach to Link Program Theory and Evaluation Data." *New Directions for Program Evaluation*. 47:93-107. 1990
- Miles, Matthew and A. Michael Huberman, *Qualitative Data Analysis: An Expanded Sourcebook*. Second Edition, SAGE Publications, Thousand Oaks, CA. 1994
- Organization of Economic Co-operation and Development (OECD). *Guidance on Evaluating Conflict Prevention and Peacebuilding Activities: working draft for application period*. DAC Network on Conflict, Peace and Development Co-operation and the DAC Network on Development Evaluation, 2008
- Patton, Michael Quinn, *Utilization-Focused Evaluation: The New Century Text*, 4th edition. Sage Publications Inc., Thousand Oaks, CA. 2008
- Patton, Michael Quinn. *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*, The Guilford Press, New York. 2010
- Patton, Michael Quinn. 2002. *Qualitative Research & Evaluation Methods*. Third Edition. Thousand Oaks; Sage Publications, Inc.
- Rogers, Mark M. "Theory-Based Evaluation," Unpublished paper submitted to the U.S. Agency for International Development (USAID/DCHA/CMM), Washington, D.C. 2010

Rogers, Mark M. "Evaluating Impact in Conflict Prevention and Peacebuilding Programs," RPP, Cambridge, forthcoming in 2012

Rogers, Patricia. "Theory-based Evaluation: Reflections Ten Years On." *New Directions for Evaluation*. No. 114:63-67. 2007

Rogers, Patricia, Anthony Petrosino, Tracy A. Huebner, Timothy A. Hacsí. "Program Theory Evaluation: Practice, Promise and Problems." *New Directions for Evaluation*. No. 87:5-13. Fall 2000

Rogers, Patricia, *et. al.* "Impact Evaluation Guidance, section 1, NONIE Subgroup 2," 2008. http://www.worldbank.org/ieg/nonie/docs/NONIE_SG2.pdf

Ross, Marc Howard. "Action Evaluation in the Theory and Practice of Conflict Resolution," *A Journal of The Network of Peace and Conflict Studies*, Vol. 8, No. 1, May 2001

Rothman, Jay. "Action Evaluation and Conflict Resolution Training: Theory, Method and Case Study." *International Negotiation*, Vol. 2, Issue 3: 451-70. 1997

Scriven, Michael. "A Meta-evaluation Checklist," unpublished, 2010
<http://michaelscriven.info/images/EvaluatingEvals-Checklist.pdf>

Stake, Robert E. *Qualitative Research: Studying How Things Work*. The Guilford Press, New York. 2010

Stufflebeam, D. L. "Evaluation Models." *New Directions for Evaluation*, No. 89 Jossey-Bass, San Francisco. Spring 2001

Weiss, Carol H. "Theory-based Evaluation: Past, Present and Future," *New Directions for Evaluation*. No. 114:41-55. 1997

Weiss, Carol H. "Which Links in Which Theories Shall We Evaluate?" *New Directions for Evaluation*. No. 87:35-45. 2000

Yarbrough, Donald B., *et. al.* *The Program Evaluation Standards*, SAGE Publications Inc., Thousand Oaks, CA 2011