interpeace

International Peacebuilding Alliance
Alliance internationale pour la consolidation de la paix
Alianza Internacional para la Consolidación de la Paz

Impact Assessment in Post-Conflict Peacebuilding

Challenges and Future Directions

Ken Menkhaus

July 2004

Abstract

The purpose of this paper is to clarify the challenges related to the assessment of impacts of peacebuilding, notably in post-conflict situations. The first part of the paper sets the context and describes how peacebuilding, notwithstanding initial resistances, eventually became accepted into development thinking and development practices. This then generated a wide and range of peacebuilding projects and processes. After years of necessary experimentation, there is now a strong desire to assess what works and what not and why. The rather technical conventional measurement tools and criteria for evaluating project impact have proven however to be of limited use when it comes to post-conflict peacebuilding. The second part of the paper presents the catalogue of challenges that have been identified in recent years. Several of these reflect generic social science problems of measurement: the search for meaningful indicators, the problem of causality and attribution, the related problem of the counter-factual and the false debate about evaluation as 'art' versus 'science'. Theoretical-conceptual problems specific to peacebuilding relate to the difficulty of separating the validity of the underlying theory of change on which the project is based from the actual implementation of the project, and from the fact that normative values and definitions of peacebuilding can be and are contested. Then there are practical obstacles such as the availability or not of reliable data, the costs of certain methods that can generate data, and general resistances to evaluation. The final section of the paper argues for practical ways forward. These can pragmatically combine qualitative and quantitative approaches, and generic with context-specific including subjective indicators, to create a 'good enough' credible narrative with supporting evidence. They also need to be attentive to unplanned for and unforeseen effects, influences and impacts. More attention needs to be paid to the underlying theories of change and how projects and processes influence the intervening variables – without being able to or required to guarantee 'Impact' especially on 'Peace Writ Large'. Finally there seems to be much potential in more regular periodic reviews and real-time evaluations, and in a judicious mix in these exercises of 'outsiders' (with context knowledge) and 'insiders' to the project.

I. Introduction: The Gradual Acceptance of 'Peacebuilding'

Interest in developing more accurate methods to measure the impact of post-conflict peacebuilding projects has grown considerably in recent years. This is the result of the confluence of two distinct trends. First is the rise of much more systematic approaches to project monitoring and evaluation in non-profit and UN development programmes in general. Prior to the 1990s, project monitoring, reporting, and evaluation tended to be relatively unstandardized and impressionistic. Over the past decade, more rigorous templates for setting defined objectives and more systematic, results-oriented measurement tools have been adapted (largely from the private sector), and are now a ubiquitous feature of virtually all work conducted by non-profits and multilateral agencies. This trend is as apparent in the large non-profit sector working on domestic assistance and outreach as it is in international aid programmes. Wherever one turns, agencies, donors, and foundations are devoting 'more' time and energy to developing tools and approaches for impact assessment. This trend is driven by a number of factors, including donor demand for greater accountability and cost-effectiveness, and a desire to capture lessons learned in a manner that fosters cumulative knowledge about various aid interventions. (OECD 2001a)

Second, the 1990s also witnessed a dramatic increase in the number of complex emergencies, protracted wars, and post-conflict reconstruction challenges – Sudan, Somalia, Congo, Rwanda, Haiti, Mozambique, Afghanistan, and Guatemala, to name only a few. These crises – and the sometimes spectacular failures of international response to those crises -- produced a growing realization that conventional development approaches were insufficient and in some instances counter-productive in post-conflict settings. This in turn led to calls for post-conflict approaches which address root causes of conflicts and which integrate peacebuilding, conflict prevention, and conflict management capacities into reconstruction and development. 'The rebuilding of war-torn societies calls for an entirely new approach to the theory and practice of development," observed Dharam Ghai, the Director of the UN Research Institute for Social Development in 1994 (quoted in WSP 1999b: 11). Over the past decade it has become clear that the implications for development assistance are staggering. A peacebuilding agenda is inherently expansive, encompassing a wide range of issues, including promotion of good governance and capacity-building as a means of conflict prevention. "Helping strengthen the capacity of a society to manage conflict without violence must be seen as a foundation for sustainable development," concluded the Development Assistance Committee (DAC) of the OECD (quoted in Rubin 2002:171).

The result has been a much greater emphasis on peacebuilding as a theme and impact in post-conflict assistance since the mid-1990s. Peacebuilding has now been "mainstreamed" into projects on education, health care, and other conventional, functional sectors. Nearly all major donors have initiated significant policy reforms to this end. (see OECD 2001b) "Donor governments, multilateral institutions and NGOs have accepted that violent conflict in developing countries undermines efforts to foster sustainable development," concludes a Saferworld/International Alert report. "In response, they have embarked on initiatives to

integrate conflict prevention into the range of development assistance instruments. The initiatives have resulted in both institutional changes and in the development of new policy and planning instruments" (Gaigals 2001: 8).

Initially, efforts to integrate peacebuilding into post-conflict assistance challenged orthodox approaches to foreign aid and met with predictable institutional resistance. Part of this resistance was bureaucratic – aid agency personnel were naturally disinclined to revise standard operating procedures. The peacebuilding approach met with hostility in some quarters because by definition it constituted a critique of conventional post-conflict assistance. Many aid agencies were particularly reluctant to accept the implicit critique that conventional aid was flawed and even doing active harm by creating conflicts and disempowering local communities. Some of the more cynical dismissed the entire concept as the latest fad in development, a new bottle into which to pour old wine.

 But the most persistent resistance to peacebuilding approaches came from aid agency officials –especially but not only from those with humanitarian missions- who saw grave danger in infusing any political objectives into their strictly "neutral" and apolitical sectoral work in potable water, primary health, education, and so on. Integrating an overt political goal into their reconstruction and development work appeared to violate a cardinal rule of foreign assistance among those many aid agency personnel who viewed their work in strictly technical terms. It ran the risk of creating political enemies and creating security risks for project staff; it forced project staff members out of their narrow sectoral expertise and into a domain in which many were acutely uncomfortable and poorly-trained; and, to the extent that peacebuilding models tend to encourage more local ownership of and extensive local participation in development projects, the additional peacebuilding objective meant almost certain reductions in perceived project efficiency. Project officials attuned to other yardsticks of success, such as execution of project goals in a timely and cost-effective manner, saw the much more participatory and consultative methods associated with peacebuilding as an unwanted and inefficient burden. Nicole Ball was one of the first to recognize this sentiment in the corridors of aid agencies. "It is important to recognize that peace building is not a distraction from development efforts," she countered, "but a critical condition for development in post-conflict environments"(1996:607).

In response to this initial resistance, the energies of proponents of peacebuilding in the early to mid 1990s were mainly devoted to advocacy. To make their case, advocates of peacebuilding focused attention on research and analysis, documenting the failures of conventional approaches in post-conflict settings, challenging the assumption that aid in conflict settings can ever be apolitical, and developing new models for post-conflict assistance aimed at promoting reconciliation. New think tanks and advocacy groups on conflict prevention and peacebuilding sprung up, and large multi-year research projects on peacebuilding and prevention were launched, from the Carnegie Commission on Preventing Deadly Conflict to the experimental fieldwork of the War-Torn Societies Project itself. Some of the critiques of aid in conflict and post-conflict settings were harsh and dismissed by aid agencies as polemic, but others struck a tone of thoughtful, constructive criticism that gradually earned a hearing among donors and aid agencies. Mary B. Anderson's *Do No Harm* was a seminal work in this regard (1999). Equally as important was the growing sophistication of the media, which by 1994 had fully absorbed the lessons of the Goma

refugee camps and had a keener eye for the many ways that presumably "neutral" and apolitical aid could actually become part of a war economy. Gradually the notion of relief and development as an apolitical humanitarian exercise in conflict zones was worn down and discredited. This was combined with vigorous lobbying of both donors and implementing agencies to convince them that post-conflict assistance strategies which focus strictly on narrow sectoral objectives are inappropriate.

The advocacy paid off. There are still many critics and skeptics of peacebuilding approaches to post-war assistance,[1] still NGO officials whose "mandate blinders" limits their work to narrow sectoral objectives, and no shortage of aid agencies which are inclined to pay lip service to the concept while pursuing business as usual in their fieldwork (Anderson 1996: 348-9). But most donor agencies, UN agencies, and NGOs have embraced the idea that post-conflict situations are very distinct from conventional development settings and hence require a separate strategy, one that builds peacebuilding objectives into post-war assistance. Post-conflict assistance is now assumed to have an essentially political purpose – a position which would have been considered radical a decade ago.

Having won the battle to integrate peacebuilding into post-conflict assistance, energies in recent years have since been devoted to project design and implementation. This has included the development of peace and conflict impact assessments (PCIA) in the planning and execution of "traditional" development projects; a considerable increase in donor interest in and funding for peacebuilding; and the rise of NGOs which focus exclusively on aspects of peacebuilding as their goal, either in advocacy work, research, or direct peacebuilding projects in conflict and post-conflict settings. Two very different types of peacebuilding projects have emerged: (1) "indirect peacebuilding," in which peacebuilding objectives are integrated into sectoral aid projects, and (2) "direct peacebuilding," projects devoted exclusively to the promotion of dialogue, consolidation of peace, building of local capacity to manage conflicts, and prevention of recurrence of conflict.

This "first generation" of peacebuilding projects is now producing a rich set of case studies, as projects either are completed or move into advanced stages. This has brought the issue of evaluation and impact assessment to the fore. Because post conflict peacebuilding is a new – even experimental – approach, project evaluation and impact assessments take on added significance, not merely as a tool insuring accountability to donors and as a mechanism with which to derive lessons learned, but as a means of testing the credibility of the entire approach. Advocates initially won the battle to incorporate peacebuilding into post-conflict development aid on the basis of very convincing, but essentially deductive, arguments. As an experimental approach, there was an element of a "leap of faith" in the claim that post-conflict development projects can and should build peace in war-torn countries. Now those arguments can and must be put to the test with actual evidence generated by recent experience.

---

[1] Critics of peacebuilding populate a wide range of the political spectrum. Skepticism about peacebuilding is most acute among so-called "realists" who consider the practice synonymous with nation-building and who view peacebuilding as naïve or worse (Betts 1996). Other analysts who express discomfort with the move to "take humanitarianism beyond the role for which it is suited" argue from a neo-idealist tradition (Rieff 2002: 112). Analysts studying interventions in "new wars" from a political economy perspective criticize what they view as the flawed neo-liberal assumptions built into much post-conflict assistance (Le Sage 1998).

But just as conventional development approaches proved inappropriate in post-conflict settings, so too have conventional measurement tools for evaluating project impact proven to be of limited use when it comes to post-conflict peacebuilding. A recent INCORE study of evaluation of conflict resolution interventions underscores the point:

> Unfortunately, evaluation theory specific to conflict resolution has not kept up with the demand, leaving the field comparatively lagging in this endeavour. As a result those engaged in peace work are seeking to meet the aforementioned needs with inadequate, and sometimes flawed, approaches and models. . . . [I]t is commonly felt that these borrowed models do not correspond to the unique circumstances found in conflict situations (Church and Shouldice 2002: 5).

The new challenge facing peacebuilding is thus that of impact assessment. How do we know when a peacebuilding project has actually built peace? What indicators are most appropriate? What evaluation tools are most useful in assessing the theories of peacebuilding on which projects are based? This is a question of vital importance to local communities engaged in peacebuilding, to aid agencies implementing such projects, and to donor agencies seeking to assess the effectiveness of funding devoted to the issue. If impact cannot be assessed, then the legitimacy of the entire enterprise of peacebuilding is open to question. Moreover, critically important opportunities for learning lessons in peacebuilding are lost if we cannot determine project impact. Without the analytic tools to determine the impact of a project on the structures and processes that strengthen prospects for peaceful coexistence, "we can only hope to list, assert, or guess at the positive or negative impacts of our actions" (Bush 1998: 4).

In the past few years, a growing number of studies and workshops have been devoted to this issue. Donor agencies, private consulting firms, academics, and peace and conflict resolution organizations are all considering the question. This paper considers the challenges identified and the emerging consensus on how to move forward.

II. Methodological Challenges to Impact Assessment

The problems associated with measuring impact of peacebuilding projects are daunting. Many reflect generic, endemic social science problems of measurement and attribution of causality. Others are more specific to peacebuilding as a project impact. The growing literature on both peacebuilding and impact assessment of peacebuilding projects has done an excellent job of cataloging these challenges. They include:

Problems of measurement. What indicators do we have to measure the extent to which peace has been built in a country or community? What indicators are most appropriate for measuring more immediate impact of a peacebuilding project on the participating community? Unlike some traditional development programs in fields such as public health, where macro-indicators of improved health are amenable to fairly precise measurement (life expectancy, infant mortality rates, malnutrition levels, etc.), or when project objectives can be easily measured and identified in advance with micro-indicators (number of children immunized in a "grab and jab" project), the objectives of peacebuilding projects are not so easily measured at either the macro- or micro-level. Peace, communal trust, and good governance are intangibles; they are process-oriented

objectives, not products. And some crucial aspects of peacebuilding are essentially psychological, involving public perceptions, not actual circumstances. Local people therefore may use very different indicators than outsiders. Efforts to operationalize those concepts by identifying indicators which are amenable to observable, replicable, verifiable measurement are famously problematic. One recent survey of the field captured the dilemma by reprinting a sobering quote from Albert Einstein in its preface: "Not everything that can be counted counts, and not everything that counts can be counted" (Church and Shouldice 2002).

Inasmuch as peacebuilding is aimed at preventing recurrence of conflict, efforts to measure project impact are also confounded by the fact that the "dependent variable" involves something that – if the project is successful -- does not happen. It is extremely difficult to take credit for things that do not occur, a fact which plagues preventive policies in many fields, from conflict management to counter-terrorism to public health. This can also adversely affect commitment to peacebuilding by project staff and donors. "Reward systems in governmental and international agencies are not designed to recognize 'nonevents,'" notes one study of prevention. "Indeed, officials drawn into preventive actions often believe they are 'being set up for failure'" (ICISS 2001: 37).

Finally, assessment of peacebuilding in post-conflict settings is constrained by "counterfactual" reasoning – namely, the impossibility of predicting what would have happened in the absence of the peacebuilding intervention. Evaluators are consistently forced to compare that which happened with that which would have happened, an exercise that is essentially speculative, not empirical, in nature. The limits of counterfactual reasoning pose a frustrating constraint on individual project evaluations, forcing such analyses to be very cautious about asserting causal links.
Notably, this is less of a problem for analyses seeking to produce broad lessons identified (whether they are 'learned' is another matter) from multiple cases. There, the focused comparison approach is a methodology which, if carefully employed, allows researchers to discern patterns and make "contingent predictions" about the effectiveness of certain peacebuilding interventions (Schnabel 2001:197).

Despite these problems, however, we know that some post-conflict situations feature greater levels of consolidation of peace than others. Identifying the most reasonable and accurate measurement of peacebuilding is thus an important prerequisite to systematic impact assessment. This is not a unique constraint – indeed, many of the most important political and social concepts (democracy, civil society) are simultaneously vital and yet difficult to measure. Over time, social scientists have made some modest progress in refining, defining, and measuring democracy and civil society; though the state of the art is far from perfect, it is increasingly close to meeting the "good enough" criterion. There is no reason to believe that peacebuilding indicators could not eventually also be measured up to a "good enough" standard as well.

Some donor agencies have attempted to draw up basic criteria for evaluations, focusing on "observable, verifiable indicators" -- impact and coverage; relevance and appropriateness; effectiveness and efficiency; timeliness; sustainability; and coherence, coordination, and complementarities (Hoffman 2002: II). There have been efforts to take this a step further by identifying a template of specific, universal, measurable indicators for peacebuilding and

good governance. A USAID paper, for instance, suggests a template of very specific and measurable indicators of consolidation of peace involving, among other things, the number of civil society groups created; likewise, the World Bank has explored very specific indicators for concepts such as governance, including such intriguing indicators as "the length of time it takes to have a telephone line installed"(cited in Bush 2002). The general consensus in the field is that peacebuilding work is not amenable to such fixed templates of indicators (see below).

Data collection. Even if good indicators can be found to measure peacebuilding impact, the specific data needed to measure those indicators can be very difficult to secure in a post-conflict setting. One could, for instance, choose hate crimes and rates of ethnic intermarriage as two plausible indicators of levels of ethnic tension and co-existence. But in the absence of standardized and reliable police reports, data on changes in the number of hate crimes committed per year is beyond reach, and no public or religious records are kept of intermarriage, that too cannot be tracked over time.   The poorer and more disrupted the society, the less likely reliable data will be available for such monitoring.

Scarce resources. Efforts to measure and assess impact, whether as periodic internal exercises or as part of external evaluations, are limited by time and resource constraints. Exercises which might go further in measuring consolidation of peace or levels of social tolerance – executed by means of household surveys or public opinion polls -- would require lengthy, extensive and longitudinal studies that they could easily exceed the cost of the actual project.   Likewise, procedures calling for an evaluator to be involved throughout the entire life of the project -- rather than simply parachuting in at the end of the project – are ideal, but require a commitment of resources to evaluation which can be hard to justify when set against pressing operational needs and limited funding. Feasibility is thus an additional constraint on impact assessment (Church and Shouldice 2002: 17).

Normative values and definitions of peacebuilding.   Related to measurement problems is the challenge of analyzing the value of the core concepts at the heart of peacebuilding projects – peace, reconciliation, conflict, justice, good governance. These are subjective notions with multiple layers of meaning and competing definitions. Reconciliation, for instance, is a catch-all term for situations ranging from simple co-existence to "liberal social solidarity" (Crocker 1999: 8). Evaluations which do not succeed in insuring that all participants share the same meaning of concepts like peacebuilding, conflict, and good governance run the risk of talking past rather than with respondents. Moreover, these values can and do exist in tension with one another, requiring local and external actors to weigh peacebuilding – which places an emphasis on consensus and reconciliation -- against approaches which emphasize justice and accountability.

Causality or attribution. Even if we develop a means of measuring the extent to which communities are more or less peaceful, we must then consider the question of a causal relationship between changes in the community and the peacebuilding project. Consolidation of peace and deterioration of peace are the results of a wide range of factors. It is virtually impossible to attribute a causal link between a change in the social or political environment and a single project, or even a bundle of projects within an aid program.   There are many other related problems of causality, such as correlations without casual links, delayed causality (in which the full impact of a project cannot be

known for many years), mutual causality, and interactive causality, to name but a few. Delayed causality is of special relevance to peacebuilding, which is generally viewed by its practitioners and advocates as a process of sowing seeds, the fruits of which may not be harvested for years to come. Insistence on evaluation mechanisms which measure impact occurring within the very short-term confines of a project cycle may fail to capture the most important effects of a project.

Conversely, a peacebuilding project could do all the right things but not produce any positive impact due to *force majeure*. Renewed conflict may break out, or communal tensions may be immune to negotiation, not because of the failure of the peacebuilding efforts, but because a powerful neighboring state intervenes or the economy collapses. Peacebuilding projects have no control over exogenous factors and cannot be held accountable when the latter produce setbacks.

For individual evaluators charged with impact assessment of peacebuilding projects, the commitment to sound social science methodology and basic epistemological principles produces a real reluctance to posit a causal link – either positive or negative – between a peacebuilding project and subsequent changes in the local political environment. Respect for social science methodologies thus stands in sharp tension with demands for impact assessment with direct causal inferences.

This latter point is important, because it exposes a false choice that has obscured much of the current discussion about impact assessment in peacebuilding. That debate has tended to frame evaluation options as a question of "art" versus "science." Those promoting a scientific approach to impact assessment portray conventional, essentially qualitative evaluation methods (typically involving an external evaluator conducting interviews, reviewing documentation, and drawing conclusions) as sloppy, anecdotal, unsystematic, and easily biased. They view the promotion of more systematic evaluation templates involving "measurable outputs" and "results-oriented" impacts as bringing more empirical rigor and comparability into the process. Critics contend that assessments which emphasize measurable outputs obscure more than they reveal by forcing evaluators to adopt flawed indicators, to ignore important intangibles because they cannot be measured, and to reduce complex social phenomena to tidy but irrelevant variables in a fool's quest to measure the unmeasurable.

In reality, both approaches to evaluation can produce insightful and accurate assessments, when properly conducted; likewise, in the wrong hands, both can produce dreadfully erroneous and useless assessments. Good social science is not only a matter of making evaluations as rigorous and systematic as possible; it is also a matter of discerning when qualitative approaches are more appropriate, when the sheer complexity of post-conflict settings demands a certain wisdom in refraining from making causal assertions. Good social science recognizes that the basic precepts of "chaos theory" in the natural sciences – in which seemingly inconsequential factors can, under the right conditions, have a greatly magnified effect (the proverbial fluttering of a butterfly's wings in China causing a storm in the Caribbean) -- are equally relevant to the chaos of state collapse and post-conflict reconstruction.

Advocates of more systematic evaluation procedures who do not take into account the serious methodological limitations of measuring impact in such projects may see themselves as disciples of scientific method, but are in reality guilty of scientism -- bad science. Indeed, nothing could set accurate impact assessment back further than the imposition of a dogmatic, mechanistic, "bean-counter" approach disguised as empiricism. For their part, skeptics of more systematic evaluation methods must not invoke the well-known problems of empirical measurement to justify sloppy thinking and resist efforts to improve evaluation. The difficulty of identifying precise indicators of impact, this argument goes, cannot be used as an excuse to forego impact assessment altogether.

The real choice is not so much between empirical versus non-empirical methodologies as it is between thoughtful, rigorous, and pragmatic approaches to project evaluation versus simple-minded, bureaucratic, and dogmatic techniques. A pragmatic approach to impact assessment is one which employs whatever evaluation tools work best; presumably this would result in creative integration of a variety of techniques ranging from qualitative interviews to more empirical measurements. In this sense, peacebuilding impact assessment is not dissimilar to good detective work, which integrates a wide range of techniques – interviews of witnesses, observation, theory-based deductions, and scientific examination of evidence – to piece together a credible narrative that seeks to approximate the truth.

Theory versus implementation. All peacebuilding interventions are based on theories about the nature of conflicts and the nature of sustained peace. In a project evaluation, it is very difficult to disaggregate assessment of the validity of the theory on which the project is based and the implementation of the project. This makes it difficult to provide an answer to the question "why?" when projects fall short of objectives. Was the implementation of the project flawed?" Or was the theory on which the project was based to blame?

Project Staff "Buy-In." One of the most immediate, practical obstacles facing peacebuilding impact assessment is the high level of suspicion, resistance, and anxiety which evaluation processes appear to inspire in project staff, especially those working at the field level. There are several reasons for this. One is a universal factor – to the extent that evaluations are viewed as mechanisms of oversight, control, and accountability by donors and headquarters, project officers will always feel a certain unease at the prospect of project weaknesses and failures bringing brought to light. This unease can take on a cultural and political dimension, particularly when most evaluations are conducted by and for 'northern' donors on work implemented by 'southern' project staff. This can produce resentment against what one researcher called the "mechanistic Northern-led quest for mainstreamable products'" (Bush 2002). Another study suggests that staff ambivalence about evaluations stems from unease with overt discussion of power in relationships between project stakeholders; the peacebuilding and conflict resolution field is in general uncomfortable with power, and prefers discussion framed by notions of partnership and collaboration. Project evaluations tend to expose the power relationships – between donor and agency, headquarters and field, project and local community – which are unavoidably part of the process (Church and Shouldice 2003:7).

Second, project teams in the field may not feel that they are stakeholders in evaluations which are mainly designed to identify lessons for application elsewhere. National staff in particular may have an exceptionally strong commitment to their own country's peace process, but view attempts to harvest lessons from their experience as an imposition and a diversion, reflecting an agenda of more importance to the headquarters team than to the field project.

Third, evaluations are resisted because they have the potential to reveal that the peacebuilding work is not succeeding, uncovering findings that "explode myths central to the belief systems of the field" and "potentially undermine the credibility of the entire field" (Church and Shouldice 2003: 6). Workshops and conferences which have brought together representatives of project staff, headquarters personnel, and donors in discussions about peacebuilding impact assessment consistently provoke such fears. This is particularly true inasmuch as the experimental nature of peacebuilding places the "burden of proof" on projects to demonstrate clear and convincing evidence of success. This is in many ways unfair – the "acceptance" of known fairly high failure rates suggests that the same burden of proof is not imposed on conventional development assistance, and that such findings do not appear likely to "explode myths central to the belief system" of major governmental and intergovernmental development assistance organizations. The historical track record of international diplomacy is also not an unqualified series of success stories. But peacebuilding approaches do not enjoy the same constituencies and are not likely to be given the benefit of the doubt if and when comparable failure rates appear in their project evaluations.

Unless project staff embrace and support impact assessment efforts, evaluations are likely to miss important observations. Creating and sustaining in the project team a strong sense that they are stakeholders in project evaluation is a crucial precondition for success.

"Teaching to the Test." One of the perennial dangers of establishing fixed benchmarks of progress is that the mechanisms for evaluating progress become an end in themselves. This is particularly likely when reward systems for personnel are too closely linked to meeting the fixed benchmarks. A benchmark is supposed to be one indicator of a broader project objective, not the objective itself. In educational circles, standardized test scores are used to measure student progress. When those scores are given too much weight in determining bonuses for teachers, the teachers "teach to the test" – devoting all of their time to prepping students to do well on the kinds of questions the standardized test will ask, thereby reducing progress in students' education to their test scores. A more pernicious example is the recent rash of scandals in American corporations, driven by an obsession with one indicator of success – stock value. By reducing the company's entire performance to a single overriding indicator, CEOs were tempted into practices and policies which yielded higher stock value in the short-term but which destroyed the company in the longer run. Impact assessment systems for peacebuilding must take care not to allow the indicators of progress to drive the project.

III. Moving Ahead

The literature on peacebuilding assessment is young – the first significant discussions of measuring peacebuilding started only in the late 90s. Though there are dozens of reports

and studies on the topic, only a handful have broken new ground and are worth highlighting. [2] As a general rule, these studies have excelled at highlighting the difficulties and dilemmas of assessment. They have also advanced our capacity to integrate conflict and peace impact assessment in the formulation and implementation stages of conventional aid projects. But they have made less progress developing new approaches and tools of measurement for assessment of actual project impact. Some have gone so far as to conclude that the field of peacebuilding impact assessment remains in a state of "methodological anarchy" while others have suggested that the quest for such assessment capacities, as part of the "Peace and Conflict Impact Assessment" (PCIA) movement, have reached a "practical dead end" (Hoffman 2002).

Despite this pessimism, recent studies and workshops on the subject have begun to advance our knowledge and evaluation capacities in a number of ways. In some cases, these advances constitute "good practices;" in other cases, a clearer delineation of a menu of options; still others, interesting ideas that have not necessarily attracted consensus support but deserve consideration. Some of the most significant advances, representing the "state of the art," are summarized below:

(1) There is a near-universal consensus that post-conflict situations vary too much to permit a pre-determined, fixed "template" for evaluation. There is no prospect for developing a bundle of universal indicators for peacebuilding in the same manner that humanitarian agencies have done for measuring impact of emergency relief.

(2) Yet most agree that post-conflict settings are not so unique as to defy comparison and generalization, or to justify determining criteria for evaluation of impact strictly on a case-by-case basis (Bush 1998; Hoffman 2002). The key then "is to find the right balance between 'off the peg' tools that are too general and 'customised' tools that are too specific to make comparison" (Goodhand 2000). Kenneth Bush, in "A Measure of Peace," stops short of identifying specific indicators on the grounds that ultimately it is locally determined criteria which matter; he advocates a "kaleidoscopic approach" to accommodate different assessment of project impact by different audiences, and to allow different stakeholders to assert ownership over the evaluation process (Bush 1998: 21).

(3) The search for both "off the peg" and "customized" indicators requires a creative a and observant eye for "sensitive indicators" – evidence which may seem on the surface to be inconsequential but which actually is a highly accurate measurement of a broader concepts or trends.[3] Robert Putnam's recent research on the decline of American civil society, "Bowling Alone," in which he uses unexpected indicators such as the decline in the number of people who bowl in leagues versus those who bowl alone, is a case in point.   By definition, such evaluation techniques require evaluators with extensive knowledge of the society in which the peacebuilding is undertaken. The need to make sure that evaluators are expert both in the general dynamics of post-conflict peacebuilding and the specifics of the country hosting the project is routinely repeated in studies and workshops, but in practice it is not always easy to find that combination in an evaluator.

---

[2] This article was written before the results of the 'Reflecting on Peace Practice Project' were synthesized. (Anderson & Olson 2003). This project sought to identify good practices in peacebuilding, but touches on several of the issues discussed here.

[3] An example of a "sensitive indicator" – one which appears narrow on the surface but which actually reveals much about the sector as a whole – is infant mortality as a general indicator of public health.

(4) The reality of multi-causality means that conclusions reached on project impact will always be tentative and modest, not definitive (Hoffman 2002). That should not be construed as a weakness of the evaluation, nor an implicit indictment of peacebuilding approaches for failure to provide clear and convincing evidence of success; it is merely a reflection of the inherent limitations of political and social inquiry.

(5) The ideal evaluation process is one that draws creatively and effectively on both qualitative and quantitative techniques and evidence, selecting the appropriate evaluation tool to match the specific question at hand. "Single-method evaluations were not considered adequate for capturing the intensity and complexity of the work" notes a recent survey of project evaluators (Church and Shouldice 2002: 11). Evaluations that produce "thick" narrative accounts of a project are more apt to yield valuable insights and lessons learned (Hoffman 2002: 5)

(6) Though local stakeholders are capable of "spinning" their observations on a project's impact for parochial reasons, they remain the single most valuable source of insights into the impact of a peacebuilding initiative. There is no substitute, notes Mary B. Anderson, for the time-honoured, qualitative technique of careful, close interviews of local stakeholders and observers. Local communities, she argues,

> 'know' whether an aid agency's program fuel the fires of suspicion and competition, or are rather seen as fair, even-handed and inclusive. And they can provide clear indications of why they know what they know. They cite evidence of cause and effect. They have opinions on impacts. These opinions provide the best source of attribution available to aid agencies. Knowing what people are saying about a program's impacts is an exceedingly important measure of its real effect (Anderson 2002)

(7) The distinction drawn between the *theories* on which peacebuilding projects are based and the actual *implementation* of the projects is of considerable help in clarifying what is and is not to be evaluated in impact assessments. Broad – sometimes competing - theories of social change, conflict, and peace inform the specific objectives of peacebuilding projects (Ross 2000: 3). For instance, peacebuilding projects that focus on peace education in classrooms are informed by broader theories about education and conflict. Projects designed to demobilize and train militiamen in peacetime occupations are informed by economic theories of peace consolidation. Peacebuilding projects that seek to use the revival of a school to increase cooperation by parents and teachers across conflict lines are informed by theories of interest group politics. And peacebuilding projects which aim to promote community dialogue and understanding across conflict lines are founded on political psychology theories of perception and misperception. Each of these theories identifies an "intervening" variable (education, economic opportunity, misperceptions, shared interests) that is believed to impact the consolidation of peace in a society. The objective of peacebuilding projects is to have an impact on one or more of these intervening variables; the theories enable us to deduce that the impact on that variable will, *ceteris paribus*, make a positive contribution to consolidation of peace. Evaluations should therefore focus on changes that the project has produced in the intervening variable as the yardstick for measuring project success. The project is not and cannot be

held accountable for changes at the "macro" level (the level of peace in the country), only for changes in the particular agent which theories tell us ought to contribute to consolidation of peace in a post-conflict setting. The task of testing and revisiting the theories themselves is an entirely separate endeavor, well beyond the scope of the project evaluation. Implicit in this "theory-based" approach to peacebuilding is the presumption that someone – academics, funding agencies, think-tanks, or others – are regularly revisiting theories of change and peacebuilding, engaging in comparative studies to test their accuracy, and routinely disseminating results of such reviews to development agencies. If that is not done, peacebuilding evaluations runs the risk of reaching conclusions which amount to "the operation was a success, the patient died" – in other words, a finding that the project achieved its narrow objectives, but that the entire exercise was built on a false premise.

(8) If "measurable outputs" are indicators linked to specific intervening variables, then it is imperative that peacebuilding project proposals identify those indicators at the outset. Broad, abstract project aims – "improving inter-community relations" – set the bar too high and distract evaluations from the more discrete and tangible impact which the project may have on intervening variables (Church and Shouldice 2003: 6). Projects must clearly identify at the outset the specific objectives they are supposed to achieve and the indicators they will use to monitor progress toward those goals.

(9) A number of initiatives have sought to establish a basic inventory of specific types of project impact. These handbooks have not produced a consensus on peacebuilding indicators, but have helped to better organize and clarify our thinking about peacebuilding impact assessment. Some, such as Bush's "A Measure of Peace" (1998), are quite broad. Bush proposes five general types of potential impact: (1) institutional capacity to manage/resolve violent conflict and to promote tolerance and build peace; (2) military and human security; (3) political structures and processes; (4) economic structures and processes; and (5) social reconstruction and empowerment
Within each of these areas he lists up to a dozen specific indicators and examples, along with sample questions an assessment might use to explore these impact areas.

(10) Project design must be informed by accurate analysis of the conflict and post-conflict environment. Efforts to develop a Peace and Conflict Impact Assessment (PCIA) methodology stress the need to begin such exercises with a strategic conflict assessment (Hoffman 2002: IV; INTRAC 1999: 6).

(11) It is universally recognized that much of the impact that peacebuilding projects have is not easily anticipated in advance, and that evaluation mechanisms must be designed to insure that the unplanned and unexpected impacts of a project, the "value added" of project work, is not overlooked. Much has been made in recent years of the need to maximize flexible response in post-conflict assistance, in order to take advantage of sudden opportunities and react to unexpected needs in a timely manner (UNDP 2002). Efforts to reform post-conflict assistance programs to that end will be fruitless if evaluation mechanisms are not also redesigned to recognize and reward flexible response and unforeseen impact. This does not preclude establishment of project objectives and indicators of success in the planning stages of a project, but suggests that those pre-set objectives constitute the point of departure for evaluation, not a fixed template.

(12) To foster greater flexibility in identifying "rolling objectives" over the course of a project life cycle, some analysts advocate an "action evaluation" model, incorporating "goal-setting, monitoring, and evaluation into a conflict resolution initiative rather than seeing these as distinct activities to be conducted independently and at different points in time" (Ross 2000: 6). This approach "embeds" both goal-setting and periodic review or real-time evaluation into the project, making it a fundamentally internal, iterative, self-reflective, and participatory approach. Because such approach makes the exercise and effort much more valuable from a (process) management point of view, it can elicit more support from field practitioners. Attitudinal changes in different quarters are required however, to make this possible. Donors must be equally committed to the approach, displaying tolerance towards shifts in objectives and willingness to cede control over those shifts to the project team and local stakeholders. The project team itself must have a sustained commitment to the approach, as it requires dedicating time, energy, and personnel to the reflective and evaluative processes. The tendency to be action-oriented, and occasionally pressures of time, funding, and staffing mean that commitments to periodic review risk being de-prioritized. And organizations whose culture tends to gravitate towards standard operating procedures may perceive periodic review processes as disruptive. The net result may be a failure to make necessary "mid-course corrections" that this approach might facilitate.

(13) Proposals to incorporate evaluation processes into the entire life of a project call into question the standard role of the individual conducting the evaluation – the external evaluator. The preference for an external evaluator – one who has no direct connection to the project – has long been justified on grounds that this better insures impartial judgment and reduces somewhat problems associated with conflict of interest. The downside of the use of external evaluators is equally obvious – the very distance which gives the evaluator his or her impartiality robs the evaluator of insight into the history and context of the project. The term which frequently depicts the arrival of an external evaluator – "parachuting" into a project – captures the essence of the problem. Alternative models, ranging from external evaluators who are detached but not uninvolved (the "family doctor" model), to a "learning facilitator" approach, to mixed evaluation teams of internal and external representatives, to self-evaluation processes (including the action evaluation model described above), are to be further tested in practice (Church and Shouldice 2003: 9-16).

Ken Menkhaus
July 2004

Bibliography

Anderson, Mary B. (2002) "Can We Know What Good We Do?" In *The Berghof Handbook for Conflict Transformation.* (http:www.berghof-center.org)

Anderson, Mary B. (1999). *Do No Harm: How Aid Can Support Peace – or War*. Boulder: Lynne Rienner Publishers.

Anderson, Mary B. (1996). "Humanitarian NGOs in Conflict Intervention." In Chester Crocker, Fen Hampson, and Pamela Aall, eds., *Managing Global Chaos.* Washington DC: US Institute of Peace, pp. 343-354.

Anderson, Mary B. & Lara Olson 2003. *Confronting War. Critical Lessons for Peace Practitioners*. Cambridge MA, The Collaborative for Development Action

Ball, Nicole. (1996). "The Challenge of Rebuilding War-Torn Societies." In Chester Crocker, Fen Hampson, and Pamela Aall, eds., *Managing Global Chaos.* Washington DC: US Institute of Peace, pp. 607-22.

Betts, Richard K. (1996). "The Delusion of Impartial Intervention." In Chester Crocker, Fen Hampson, and Pamela Aall, eds., *Managing Global Chaos.* Washington DC: US Institute of Peace, pp. 333-342.

Bush, Kenneth. (1998). "A Measure of Peace: Peace and Conflict Impact Assessment (PCIA) of Development Projects in Conflict Zones." Working Paper #1, IDRC, The Peacebuilding and Reconstruction Program Initiative, Ottawa.

Bush, Kenneth. (1995) "Fitting the Pieces Together: Canadian Contributions to the Challenge of Rebuilding Wartorn Societies." IDRC, Ottawa, unpublished paper (May).

Chavez, Maria Patricia Gonzalez. (2002) "Conflict Resolution and Evaluation." Report of an INCORE Meeting, Londonderry, Northern Ireland, (June).

Clerc, Valerie. (2000). "The Rehabilitation of War-Torn Societies: A Bibliography." CASIN, Geneva.

Church, Cheyanne, and Julie Shouldice. (2003). *The Evaluation of Conflict Resolution Interventions Part II: Emerging Theory and Practice*. Londonberry, Northern Ireland: INCORE. (http://www.incore.ulst.ac.uk)

Church, Cheyanne, and Julie Shouldice. (2002) *The Evaluation of Conflict Resolution Interventions: Framing the State of Play*. Londonberry, Northern Ireland: INCORE.

CIDA. (2002) "Results Oriented Logical Framework for Peacebuilding Projects."

Collaborative for Development Action. (2002). "Effectiveness Criteria." CDA: Reflecting on Peace Practice: Issue Papers (http://www.cdainc.com)

Crocker, David A. (1999). "Reckoning with Past Wrongs: A Normative Framework." *Ethics and International Affairs*, vol. 13, accessed via http://www.ciaonet.org/lojcceia/cceia_99crd01.htm

Duffield, Mark. (1997). *Evaluating Conflict Resolution: Context. Models, and Methodology*. Norway: Chr. Michelsen Institute.

European Platform for Conflict Prevention and Transformation. (2002). *Towards Better Peace Building Practice: On Lessons Learned, Evaluation Practices and Aid and Conflict.*

FEWER/International Alert/Saferworld. (2001). "Development in Conflict: A Seven Step Tool for Planners." David Nyheim, Manuela Leonhardt, and Cynthia Gaigals.

FEWER/International Alert. (1999). "Promoting Development in Areas of Actual or Potential Conflict: Approaches in Conflict Impact Assessment and Early Warning." Manuela Leonhardt and David Nyheim.

Gaigals, Cynthia, with Manuela Leonhardt. (2001). "Conflict Sensitive Approaches to Development: A Review of Practice." Saferworld, International Alert, and IDRC.

ICISS. (2001) *The Responsibility to Protect: Research, Bibliography, Background*. Supplementary Volume to the Report of the International Commission on Intervention and State Sovereignty. Ottawa: IDRC.

INCORE. (2002). "Conflict Resolution Evaluation Meeting Session Summaries." (July 4-5).

International Alert. (1998). "Conflict Transformation Work: Code of Conduct." London, International Alert.

International Alert. (2000). "Development/Humanitarian NGOs and Conflict: A Bibliography and Listing of Web Sources." Andrew Sheriff and Eeva Vaskio, compilers.

Johannsen, Agneta. (nd) "Measuring Results of Assistance Programmes to War-Torn Societies." Geneva: UN Research Institute for Social Development.

Leonhardt, Manuela. (2001). "The Challenge of Linking Aid and Peacebuilding." In *Peacebuilding: A Field Guide* edited by Luc Reychler and Thania Paffenholz. Boulder: Lynne Rienner Publishers, pp. 238-45.

Leonhardt, Manuela (1999). "Improving Capacities and Procedures for Formulating and Implementing Effective Conflict Prevention Strategies: An Overview of Recent Donor Initiatives." London, International Alert.

Leonhardt, Manuela (2002). "Towards a Unified Methodology: Reframing PCIA." Berghof Center.

Le Sage, Andre. (1998). "Engaging the Political Economy of Conflict: Towards a Radical Humanitarianism." *Civil Wars* Vol. 1, no. 4 (Winter), pp. 27-55.

Lund, Michael (1996). *Preventing Violent Conflicts: A Strategy for Preventive Diplomacy*. Washington DC, USIP.

Lund, Michael. (2000). "Assessment Questions." Greater Horn of Africa Peacebuilding Project, USAID.

Lund, Michael, and Guenola Rasamoelina, eds. (2000). *The Impact of Conflict Prevention Policy: Cases, Measures, Assessments*. SWP-CPN. Nomos Verlagsgesellschaft, Baden-Baden Germany.

OECD 2001a: *Evaluation Feedback for Effective Learning and Accountability*. Paris, OECD Evaluation and Effectiveness Series Report no. 5

OECD 2001b: *The DAC Guidelines. Helping prevent Violent Conflict*. Paris, OECD

Reychler, Luc, and Thania Paffenholz, eds. (2001). *Peacebuilding: A Field Guide*. Boulder: Lynne Rienner Publishers.

Rieff, David (2002) "Humanitarianism in Crisis." *Foreign Affairs* vol. 81. no. 6 (November/December), pp. 111-121.

Rubin, Barnett (2002). *Blood on the Doorstep: The Politics of Preventive Action*. New York: Century Foundation Press.

Schnabel, Albrecht (2001). "One Size Fits All? Focused Comparison and Policy-Relevant Research on Violently-Divided Societies." In *Researching Violently Divided Societies: Ethical and Methodological Issues.* Edited by Marie Smyth and Gillian Robinson. New York: United Nations University Press, pp. 193-206.

SWP-CPN (1999). "Conflict Impact Assessment: A Practical Working Tool for Prioritising Development Assistance in Unstable Situations." SWP-CPN, Brussels and Munich.

UNDP Evaluation Office. (2000). "Development Effectiveness: Review of Evaluative Evidence." New York: UNDP.

USAID (2001) "Evaluation of Recent USAID Evaluation Experience." Cynthia Clapp-Wincek and Richard Blue. USAID, Center for Development Information and Evaluation Working Paper #320. http://www.dec.org/pdf_docs/PNACG632.pdf

USAID (2000). "Maximizing Impact and Sustainability." Washington, DC, USAID.

USAID (2000) "TIPS: Performance Monitoring and Evaluation." Series of reports (1996-2000).

Wood, Bernard. (2001) "Development Dimensions of Conflict Prevention and Peacebuilding." UNDP. (June).

WSP (1999a). *Rebuilding After War: Lessons from the War-Torn Societies Project.* Geneva: WSP.

WSP (1999b). *The War-Torn Societies Project: The First Four Years.* Geneva: WSP.

interpeace

7-9 chemin de Balexert
CH-1219 Châtelaine—Geneva
Switzerland

www.interpeace.org