

ASSESSING THE IMPACT OF INTERNATIONAL VOLUNTEER CO-OPERATION

GUIDING QUESTIONS AND CANADIAN EXPERIENCES

a discussion paper for IVCO 2012 by Daniel Buckles and Jacques Chevalier



Foreword

This is the eleventh in a series of discussion papers produced by the International Forum on Development Service (Forum), which follows on from our research work on trends in international volunteering and cooperation in recent years.

This paper aims to consider some of the thinking and tools for assessing impact in the international volunteering and cooperation sector, what we can learn from this, and identify some challenges for the future.

The views expressed in this paper are not necessarily those of Forum or its members or of the organisations for which the authors work. The responsibility for these views rests with the authors alone.

Nita Kapoor
Chair of Forum

About Forum

The International Forum on Development Service (known as “Forum”) is the most significant global network of International Volunteer Cooperation Organisations. Forum aims to share information, develop good practice and enhance cooperation and support between its members. Together, Forum members explore innovative practice and research key contemporary issues, focusing on organisational learning and improved practice. This information is shared in person, at conferences and via the website.

Contents

Introduction	3
What is Impact Assessment?	4
Why Do Impact Assessment?	11
Who Should Be Engaged?	12
How to Assess Impact?	14
Conclusion.....	19
References	20

Introduction

Many Civil Society Organisations (CSOs) working globally are looking for ways to assess the impact of their work. The search is prompted by a desire to learn about and improve the effectiveness of international development programming, and satisfy the demands of widely used accountability frameworks such as Results-Based Management (RBM). The government-sponsored Paris Declaration on Aid Effectiveness, and CSO responses to the agreement, highlight the challenge, which consists in striking a meaningful balance between two legitimate expectations: honest and useful learning from ongoing work, and accounting for both resources and results to donors and, importantly, to the intended beneficiaries of development.

Pressures to demonstrate that international aid is making a difference, and to learn what individual components of a program work best, are reasons why assessing impacts has become a goal for international volunteer cooperation (IVC) programs. For many decades now, CSOs have been experimenting with a range of IVC programs, including partnership models and horizontal, reciprocal volunteering models aimed at building capacity and sustainable results. Some have given attention to the involvement of volunteers from developing nations through South-to-South exchanges, while others have shifted to shorter-term educational programs for volunteers from the Global North. The impacts of these activities remain largely undocumented, however, and little is known about the different outcomes of different approaches. The few studies that have been done focus primarily on the benefits of volunteering for the volunteers (Tiessen and Heron, 2012; McBride, Lough and Sherraden, 2010), not the consequences of volunteering for other stakeholders.

Part of the problem lies in donor and CSO emphasis on accounting for the delivery of outputs (goods and services) traceable to specific activities and inputs. Christie (2008) notes that there are few incentives for CSOs to go beyond this extremely narrow accountability framework to examine impacts. Dependence on evaluation findings for the renewal of funding discourages fund recipients from reporting project weaknesses or negative outcomes and findings questioning the deeper relevance of their work. Nor does it help that there is widespread doubt concerning the extent to which donors actually share and make good use of all the reports they receive or commission.

Davies (2005) argues that reluctance among CSOs to engage with questions of impact is, in and of itself, a problem. He laments the growing tendency of organisations to limit their responsibilities to “the output level”, defining impacts so narrowly that they do not pay attention to system-wide changes and longer-term impacts, or attempts to generate knowledge about what works and does not work in a given situation. Support for greater attention to impact assessment can also be found in recommendations from formal evaluations of international volunteer cooperation programs (Jackson et al., 2007).

While there is a growing recognition of the importance of assessing impacts, how to do so remains a challenge. The range of designs and methods for impact assessments is narrow (Stern et al., 2012), and little attention has been given to adapting them to the particular circumstances of IVC programs. This paper begins to address these gaps by identifying key questions and methodological strategies organisations can use when developing impact assessments. Building on the work of Chevalier and Buckles (2012), the paper draws on recent impact assessment literature and work with a Canadian community of evaluation practice that includes seven of the nine Canadian CSOs currently receiving funds from the Canadian International Development Agency (CIDA) for IVC programs.

We argue that it is possible to develop meaningful understandings of the impact of international volunteer cooperation so long as the assessment process is creatively designed and scaled to fit the purpose and situation. To achieve this, further innovation in the development of evidence- and people-based methodologies is needed. We suggest that to assess impacts in complex settings, methods need to bring together the processes of fact-finding and the construction of meaning in complex settings currently separated in mainstream methodologies.

What is Impact Assessment?

In recent years, efforts to assess the impact of international development on people's lives have become a lion's den of competing terminologies and methodologies associated with the results agenda¹. The debate raises three central issues in the broader field of monitoring and evaluation: **time frame**, **system boundaries** and **causality**. The first issue concerns the understanding of results and the distinction between outputs, outcomes and impacts that is now part and parcel of the monitoring and evaluation (M&E) lexicon. These are typically seen as points on a timeline, from **immediate to intermediate and ultimate or final results**. Impact assessment (IA) implies responding to questions regarding results of an intervention that appear over longer time frames, often five years or more. For programs involving international volunteer cooperation, this might refer to changes in communities after many years of hosting volunteers, or changes in volunteers years after their participation in the program, as with Canada World Youth's assessment of impact over a ten year period (see Box 1).

¹ For a comprehensive and state-of-the-art review of the literature on impact evaluation in the field of international development, see Stern et al. (2012). We use the term "impact assessment" instead of "impact evaluation" in this paper to reduce the emphasis on the role of outside experts often implied by the literature on impact evaluation. Impact assessment is a term also widely used in the environmental field and in some areas of policy studies. In these contexts, it refers to studies that assess the **future consequences** of a current or proposed action, such as the potential environmental consequences of the Northern Gateway pipeline. The assessments are forward-looking, and imply public input with a view to anticipating and managing impacts, as opposed to the backward-looking focus of impact evaluation in the international development field.

Box 1: Impacts over time

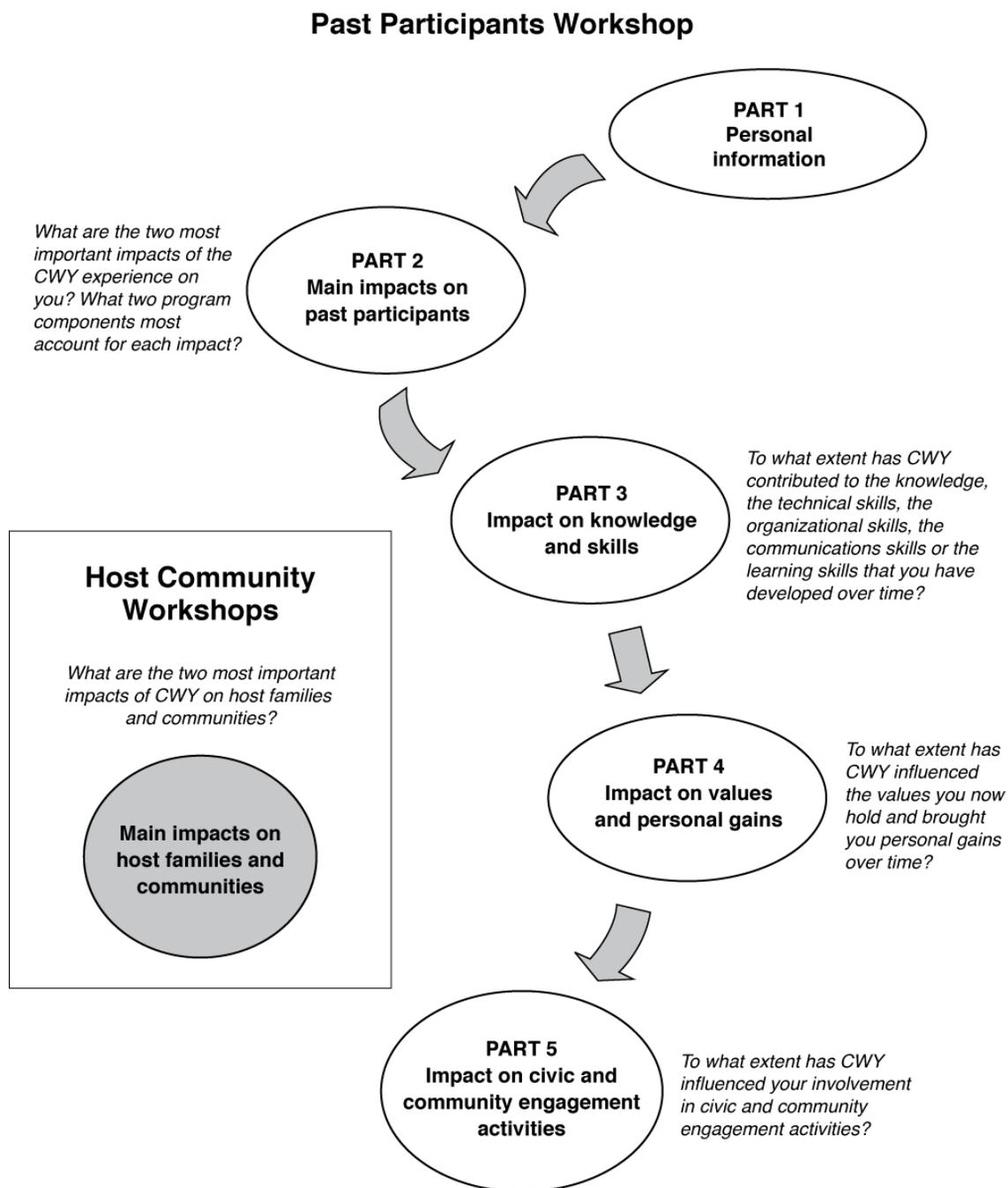
Canada World Youth (CWY) is a Canadian non-governmental organisation focused on youth programming in Canada and abroad. In recent years, they have conducted a number of impact assessments dealing with different programs and aspects of their programs.

In 2006 (McLaren and Turcot), the organisation undertook an assessment of the impact of a reciprocal youth exchange program on participants and communities in five countries, including Canada, Cuba, the Ukraine, Benin and Thailand. The assessment examined impacts from the point of view of past and present participants and communities with a view to understanding and comparing intermediate and longer-term impacts. It also examined a range of impacts from the point of view of host communities. It was a large-scale assessment involving reflections by different stakeholder groups (youth, community hosts, program staff) involved in the program over a ten-year period, with an effort to roll up analyses from several countries and continents.

The overall design, summarised in Figure 1, addressed questions related to impacts not only on knowledge and skills but also the influence of the program on the values and attitudes of participants, perceived personal and professional gains, and influence on involvement in civic and community engagement activities later in life. Importantly, individuals and community groups also provided feedback and perspectives regarding the impacts of the program on families and host communities.

The final assessment recommended that CWY *ground its current emphasis on individual learning objectives in specific community projects*. This subtle strategic shift implied adjustments to the work placement component of the exchange program, to put both *work and community experience* at the centre of the learning agenda. The anticipated effect would be to have a longer-term impact on host communities while at the same time building relevant individual technical or professional skills among volunteers, in addition to communication, learning and organisational skills.

Figure 1 Canada World Youth impact questions and evaluation design



Second, efforts to assess impact reflect a deep concern with the **large-scale, system-wide goals** that justify development initiatives in the first place — eradicating extreme poverty and hunger, achieving universal primary education, reducing child mortality rates, combating HIV/AIDS, malaria and other preventable diseases, and other major results expressed in the Millennium Development Goals (MDGs) and other development agendas. Impact assessment aims to demonstrate progress (or the absence of progress) towards the large-scale goals within which specific interventions are embedded. This means assessing changes in whole systems and across systems, not just a few elements in a narrowly defined system. For IVC programs, this might involve showing how changes in systems of knowledge and awareness across geographic and cultural boundaries help to create the conditions for change at broader levels (Mwathi and Perold, 2012).

People often assume that systems thinking must focus on the whole project or program, with a clear understanding of the corresponding goals — desired results that are coherent, shared by all parties, and stable over time. At first sight, holistic ventures seem to set us up for failure. Attempts at the comprehensive view, with plans to act on it in the fullness of time, can guide people and organisations into a morass of competing priorities, unprovable assumptions and inflexible responses to persistent messiness and uncertainty. Unrealistic plans and overly optimistic accounts of final outcomes often result.

Using a systems perspective to focus on a more specific set of activities can respond to this challenge by avoiding the pretension of holism and oversimplifications that ignore vast differences in local settings. IVC programs, for example, are embedded in long chains of expected results and associated activities. However well integrated the program may be, each set or subset of activities may have its own goals or desired results that stakeholders may assess in their own right. Ends are often embedded in the means taken to achieve them (community engagement or gender equity, for instance), with the implication that process matters as much as outcome.

Still, attempts to see the forest from the trees are essential. People and organisations are constantly struggling to understand and manage complex situations. The question about broader, distant results warns us against focusing on the local and immediate alone. Both are needed to make sense of system boundaries and interactions, and means to achieve large-scale, system-wide goals (Box 2).

Box 2: Assessing program results and interactions (using *Activity Dynamics*)

In Mozambique, Cuso International staff assessed the interaction of different activities in their program as part of a developmental evaluation undertaken in November 2011 (Owens, 2011). The purpose was to help build a theory of change with partners by assessing the dynamic links between program activities expected to contribute to “bringing people together to fight poverty”.

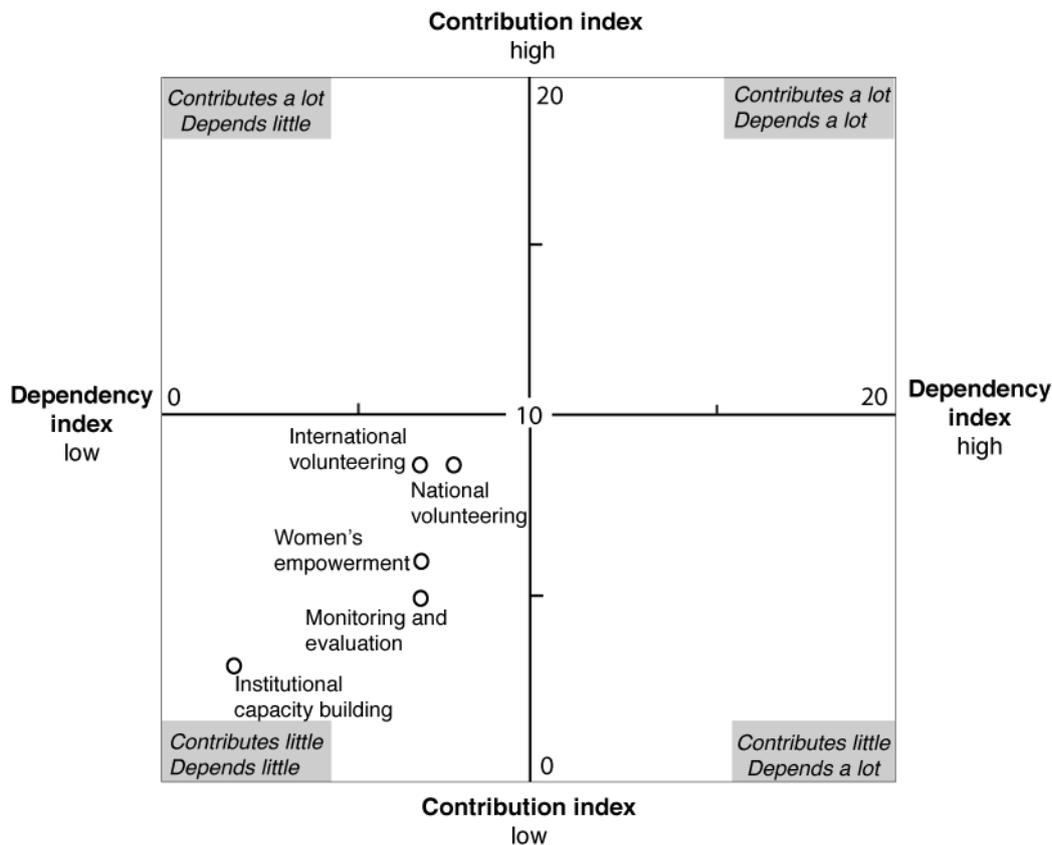
The program mobilises Canadian and Mozambican volunteers, contributes to women’s empowerment through a gender fund, monitors and evaluates program activities and builds institutional capacity through a Program Capacity Development Fund (PCDF).

The assessment, undertaken in Mozambique by various stakeholders in the program, used the *Activity Dynamics* tool (Chevalier and Buckles, 2012) to describe how activities in the program interact with each other, using a rating scale to indicate the level of effective interaction or program synergy.

The result, shown in Figure 2, shows that volunteering activities, both by Canadians and Mozambicans, contribute the most to other activities of the program (higher in the figure). On the whole, however, the interaction between program activities is weak; each activity makes a limited contribution to other activities. This means that activities and resources of the program are dispersed. While important in themselves, institutional capacity building activities operate in relative isolation from other activities (lower left quadrant).

The observations supported discussion by participants, including the Country Director, Program Managers, staff and volunteers, about how to implement activities and ways to do these differently so that they could contribute more to other parts of the program. Participants made plans to repeat the exercise after some time to see if greater synergies among activities had been achieved.

Figure 2 *The interaction of Cuso International program activities, Mozambique (using Causal Dynamics)*



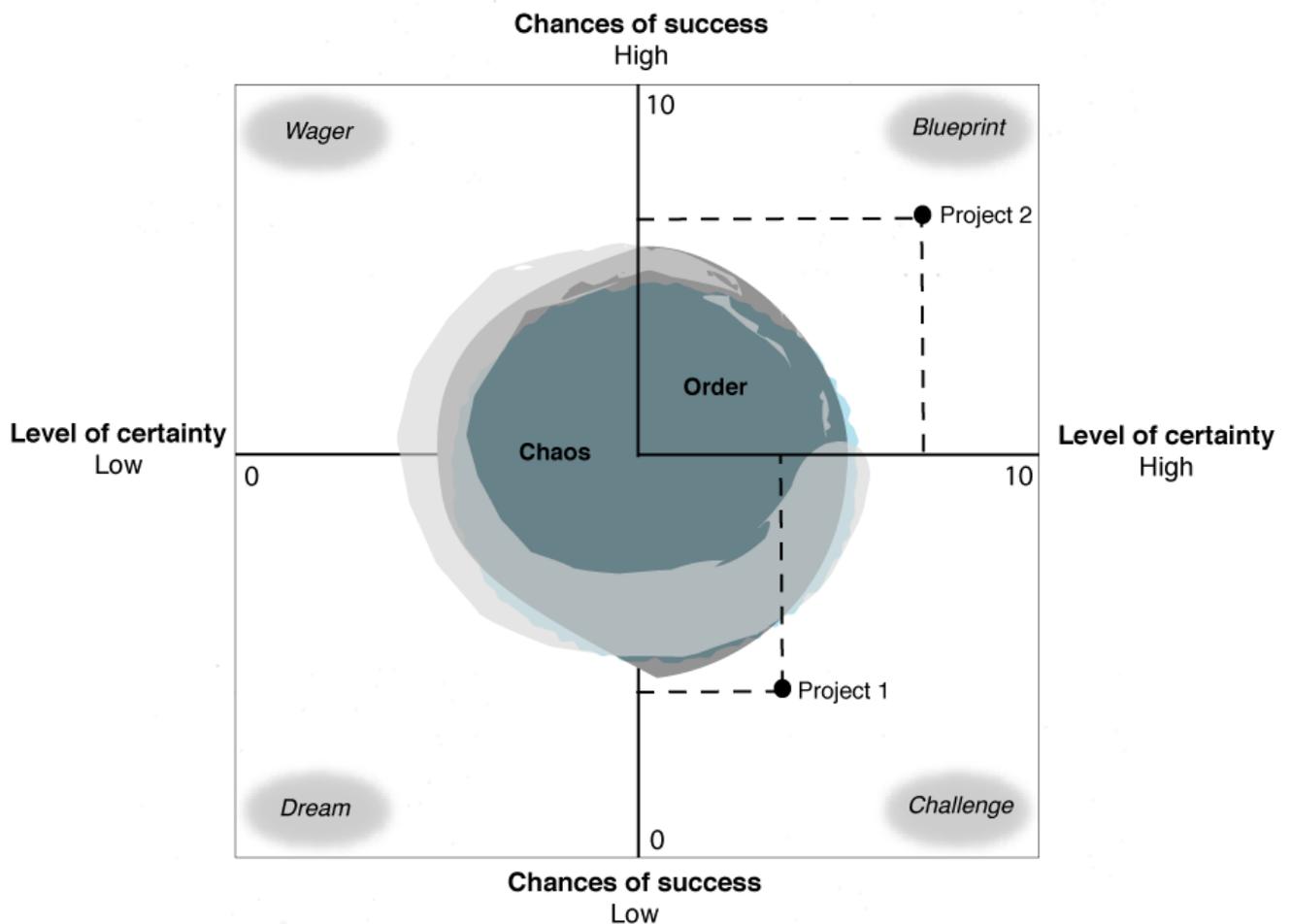
The third issue built into the IA debate concerns notions of **complex causality** – the relationship between an event (the cause) and another event (the effect). Explaining the causal links between results and the interventions to which they can be attributed is complex and can take many forms, including counterfactual thinking, the search for root causes and the examination of the underlying conditions upon which desired results rest. For IVC programs, this might mean positing plausible causal links between expected results and proposed activities based on a theory of change, and then collecting evidence of observed changes and the logical link to particular activities or components of a program.

Impact assessments run the risk of reducing causality to its simplest expression, i.e. single and linear cause-effect relationships, leaving out considerations of complexity in society and nature. While it is now commonplace to say that our world is complex to the *n*th degree, it is another thing to try and map out that complexity in a chain of causes and effects as results-based management frameworks attempt to do. Planning and reporting methods that attempt to do so, while useful under conditions that are relatively simple and orderly, are not credible ways to generate understanding of contexts where complexity and uncertainty are high. This is particularly true of modern aid programs, including most IVC programs, which emphasise capacity development, decentralised implementation and multiple partnerships where many different factors and actors interact over long periods of time.

On this matter, the tool *Order and Chaos* (Chevalier and Buckles, 2012) offers evaluators and other stakeholders a means to reflect on when linear approaches to causality make sense and when they do not. It addresses the question of the complexity of interventions by supporting thinking and dialogue around the **chances** of achieving project or program goals, and how **certain** or confident people are that the information and knowledge they have (about the conditions and factors affecting the project or program) is complete and reliable.

The four quadrants created by the graph (Figure 3) point to two scenarios based on varying levels of uncertainty and complexity. The first scenario (top left quadrant and the two lower quadrants) refers to projects or programs where information and knowledge is very incomplete, links between causes and effects are not linear and straightforward, or chains of actions, partners and results are complex. This is the world of relative “chaos”, a context that calls for an iterative and continuous approach to impact assessment guided by working hypotheses and life-world experimentation. Evaluative thinking performed in this context calls for progressive judgments in the middle of an ongoing process where general and specific objectives **interact and evolve**, subject to negotiations, compromise and change over time. As discussed further, below, organisations involved in projects of this nature would benefit from stakeholder engagement and capacity building in analytical thinking, as needed, and with the level of detail needed to inform actions and impact assessment undertaken along the way (Box 3).

Figure 3 Assessing the complexity of interventions



By contrast, the upper right quadrant points to projects or programs where the chances of success are good and much is known about the conditions needed to achieve success. They are characterised by a coherent set of objectives shared by all stakeholders that are clearly achievable with a well-defined set of inputs (time, resources, people). In this world of relative “order” and predictability, a linear logical model can guide impact assessments that:

- establish the initial situation;
- implement activities reflecting logical links to expected results;
- monitor the ongoing results of implementation against the initial set of observations;
- form final judgments of results against goals using relevant criteria and indicators.

The initial situation (baseline) can also be re-examined in hindsight, to produce effects of Socratic learning such as “Now we know what we knew” or “Now we know what we didn’t know” (Chevalier and Buckles, 2012).

Box 3: Causality in complex settings (using *Order and Chaos* and *The Socratic Wheel*)

Planning under uncertainty

The Alliance of Youth (la Segovia, Nicaragua) engages young, rural micro-entrepreneurs in activities aimed at creating new economic opportunities in a region with high levels of youth unemployment. These involve the production and marketing of coffee, honey, fruits and vegetables. A volunteer with Canada World Youth facilitated an assessment of project objectives related to honey production by youth members of the alliance using the tool *Order and Chaos*. The youth concluded that their chances of achieving income gains seemed high. However, they needed more information about factors affecting honey prices and the required hygienic standards of their products before they could proceed with further project planning. The group thus considered various actions such as additional research and training on quality control and processing methods, and broadening their contacts with buyers. Planning would also need to take into account many other factors and conditions that are likely to affect their chances of success, including honey prices, weather and access to sufficient labour in a context of high out-migration. Given that many of these factors are subject to unpredictable fluctuations, expected results and activities to achieve and evaluate them may have to be modified along the way.

Planning under favourable conditions

The Global Learner Program at Canada World Youth is a self-financing international volunteer cooperation program focused on creating an educational experience for Canadian youth aged 15 to 30, typically in partnership with high schools, colleges and CEGEPs (General and Vocational Colleges). The program involves living with a host family for a period of two to six weeks, and engagement in community projects run by host partner organisations. Given these parameters, impact assessments can be designed in advance in considerable detail, using clear methods to assess final results against initial expectations. For example, the program uses *The Socratic Wheel* to support self-assessments by youth of their level of knowledge and skills in different areas prior to departure. Program plans to help develop and evaluate the development of skills and knowledge over time are then created and implemented. Following their return, the youth assess their actual level of achievement in each impact area. They also revisit their pre-departure assessment in light of their revised understanding of themselves and their strengths and weaknesses. These estimates and the corresponding methods allow CWY to evaluate the real impact on youth skills and knowledge, attributable to the exchange experience.

Impact assessment purports to address results that occur over the long term, usually on a relatively broad scale and in complex settings. Many IA methodologies, however, are more narrowly designed and deal with results that are achieved either over time, on a systemic scale *or* in complex settings. For example, an assessment based on a Logical Framework and survey design may focus on empirical and descriptive accounts of changes over time, without getting into causal theory and explanations of multi-directional linkages. Alternatively, attention may be given, as in Outcome Mapping, to how and why outcomes have been achieved (or not achieved), without attempting to document, let alone measure, longer-term impacts. In our view, IA should be able to address longer time frames, broader system goals *and* complex causation. The weight given to each dimension of impact may vary, however, depending on the purpose of the exercise. IA should answer other critical questions as well, such as why the assessment is needed in the first place, whom it is for, who should be involved and how it should be done.

Why Do Impact Assessment?

Decisions regarding impact assessment and flexible ways to address the question of “ultimate results” are closely tied to the actual motivations that underlie any evaluation process. Why do the assessment in the first place, and for whom?

The recruitment and placement of international volunteers with local development groups, the core activity of IVC programs, involves many stakeholders affected in different ways by volunteer work and for whom an impact assessment could be designed and carried out with the corresponding goals and methods: donors, the ultimate beneficiaries of development activities, partner organisations in developing countries, the volunteer sending organisation itself and volunteers themselves. Many IVC programs involve an even more extensive chain of relationships among stakeholders, each of whom may have different information needs and different uses for a specific impact assessment. The challenge is to design an assessment that fits the overall purpose or multiple goals for which it is being undertaken in the first place.

Broadly speaking, there are three possible purposes or motivations for doing an impact assessment: to **account for resources and results**, to **learn for the purpose of planning**, and to tell stories about successes and lessons learned to **guide and inspire others** (Chevalier and Buckles, 2012). While the importance of learning for planning and sharing inspiring stories is widely recognised, the *de facto* practice of many CSOs and donors in international development is to focus on accounting for resources and results. Yet the three possible motivations are equally important and need not be mutually exclusive, depending on the motivation.

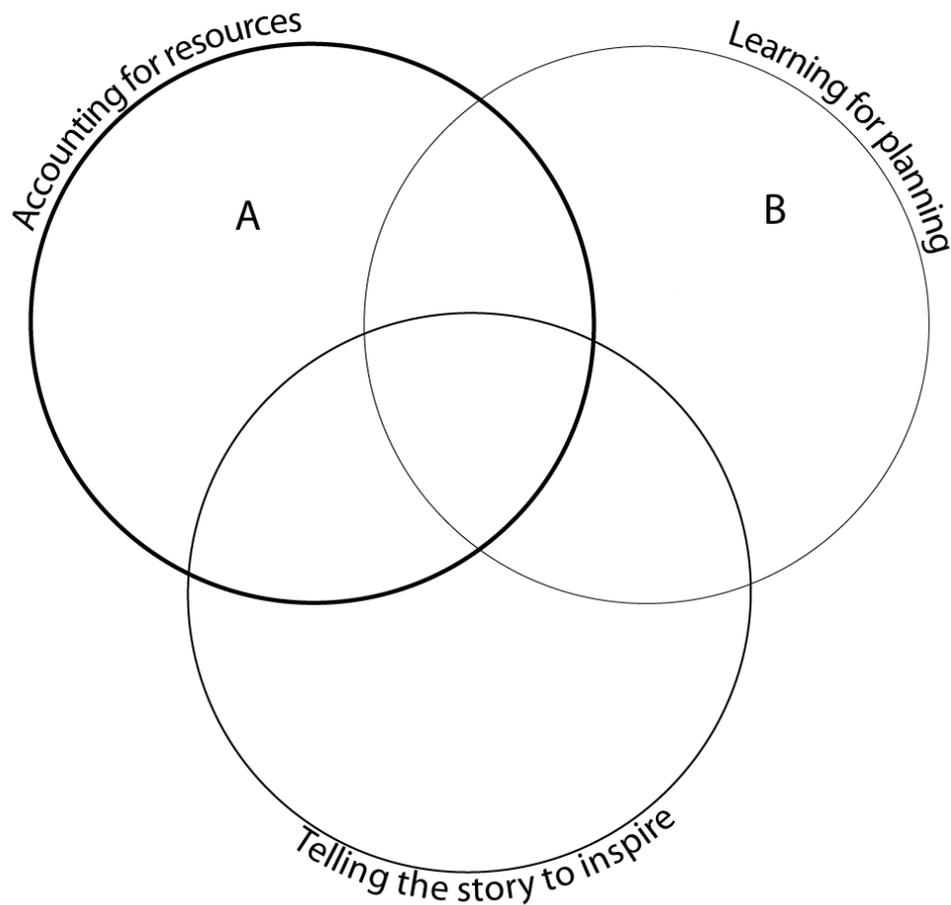
Figure 4 provides a method stakeholders can use to reflect on and negotiate around the purpose(s) of a specific assessment or a broader M&E framework. The key is to adjust the assessment process to fit the needs of stakeholders, whether it be a single purpose or a combination of several purposes. If accounting for resources and results is the sole purpose of the impact assessment, then it is likely that greater attention will be needed to collecting quantitative information on what has changed and the extent of change resulting from volunteering programs. By contrast, impact assessments aimed at inspiring others may choose to focus on collecting Most Significant Change stories or using video to capture the stories showing the difference volunteering has made in the lives of individuals and communities. Assessments with both accounting and inspirational motivations in mind cannot use a single methodology to achieve their twofold purpose. They would be well advised to use mixed methods (see Box 4).

Box 4: CESO revisits its M&E system

CESO, the Canadian Executive Service Organization, is one of Canada’s leading volunteer-based development organisations with a focus on strengthening economies through the transfer of skills and knowledge by highly skilled and experienced Canadian professionals to developing communities in Canada and around the world. A comprehensive review of the organisation’s M&E system and activities found that historically these did not direct feedback and information to key stakeholders in the process, making it difficult to plan future activities. Evaluation activities were often out of sync with the planning processes of particular stakeholders, and failed to generate the information needed to inform actions directly. The assessment also concluded that virtually all evaluation activities were aimed at accounting for resources, when stakeholders actually cared about several purposes.

The revised M&E system following the review distinguishes broadly between two evaluation frameworks: program evaluation (point A in Figure 4) and process evaluation (point B). Program evaluation is primarily for donors interested in accounting for resources and explaining variances between expected and actual results. Information acquired includes measures of organisational performance and personal performance relevant to understanding the quality of program results. The process evaluation framework, by contrast, is primarily intended to assist CESO in learning to plan improvements in the effectiveness of service delivery. Information on the factors limiting recruitment of particular kinds of volunteers, and ways to overcome these constraints, is an example of the kind of information needed for this purpose.

Figure 4 *Monitoring and Evaluation at CESO*



A thoughtful and deliberate consideration of the purpose of the assessment and knowing whom it is for opens impact assessments to the possibility of **upward and downward accountability and partner-to-partner** evaluations as well. Without careful treatment of this question, impact assessment will inevitably slide into a process of upward accounting from lower to higher levels of stakeholders in the chain, ignoring the information needs, expectations and perspectives of partner organisations and “lower level” stakeholders. Ideally, decisions regarding purpose and stakeholder goals reflect discussion and agreement between the parties concerning the best strategy to adopt.

Who Should Be Engaged?

What constitutes an impact, and explanations of the associated complex causes, long term results and system-wide goals, will vary in light of different stakeholder interests, viewpoints and theories of change. Perspectives may also vary within a stakeholder group, from one activity to another, and over time as well. Most approaches to IA, however, are expert-based, rarely engaging the stakeholders in identifying impacts or reflecting systematically on how and why things have changed. This problem applies not only to the intended beneficiaries of projects and programs but often also to the staff of the implementing agencies and the officials and general citizenry of both recipient and donor countries. These stakeholders, including the staff of IVC programs, are typically reduced to being providers of raw data. How to engage multiple stakeholders in developing and reflecting on the longer term, causal reasoning and ultimate goals embedded in IVC programs is a major challenge, addressed in the following section.

Decisions regarding who should be engaged in an impact assessment cannot be determined in the abstract. Nor can stakeholder involvement always be planned in advance. While the use of participatory self-assessment methods has greatly broadened the scope for stakeholder engagement in evaluations, it is important to note that participation does not preclude third-party expert contributions. Co-evaluations (collaborative) and exogenous, expert-based evaluations are equally valid options that can feed into each other, depending on the purpose, the time frame and resources available. Each strategy has strengths and weaknesses. More often than not, impact assessments for the purpose of accounting to donors for resources would involve the intervention of outside experts (ideally neutral in relation to all stakeholders and selected with an agreement between the parties). By contrast, impact assessment focused on learning creates good conditions for self-assessments or collaborative evaluation strategies, with or without third party input (see Box 5).

Box 5: Collaborative impact assessment (using *The Socratic Wheel*)

Canada World Youth's (CWY) international volunteer cooperation program seeks, among other things, to improve the performance of local partner organisations responsible for delivering a range of services for poor and marginalised communities. It focuses on four aspects of organisational performance: effectiveness, efficiency, relevance and financial viability. Improvements among their partners in these areas are expected to contribute to stronger organisations, and more effective ultimate outcomes.

An impact assessment process launched by CWY Program Managers in June 2011 opted to ground measurements of partner organisational performance in local objectives and indicators, rather than using a single yardstick against which the performance of all partners would be measured by outside experts. This reflected recognition that organisational performance is complex and depends on many contextual factors that vary from partner to partner. In reality, the starting point and organisational objectives of different partners, and the steps and capacities needed to make progress, vary considerably. The assessment consequently set aside direct comparisons among organisations and focused instead on the extent to which organisations actually achieve performance objectives they set for themselves, regardless of size or current capacity. The design (Figure 5) developed with CWY managers and refined further with local partners involved a series of steps over a period of 12 months.

Step 1

Partners and CWY do a stakeholder analysis to determine who should contribute to the assessment process. They jointly convene a meeting of stakeholders to co-construct a common understanding of four areas of organisational performance and set specific objectives for the organisation related to each performance area.

Step 2

Participants create a set of progress markers (statements of graduated change) for each performance objective that

would represent observable change in the way the organisation works over a 12-month period. These statements represent levels of achievement of the specific performance objective, from achievements the organisation 'expects to see' to those they would 'like to see' and those they would 'love to see' (see Earl, Carden and Smutylo, 2001 for a full description of progress markers).

Step 3

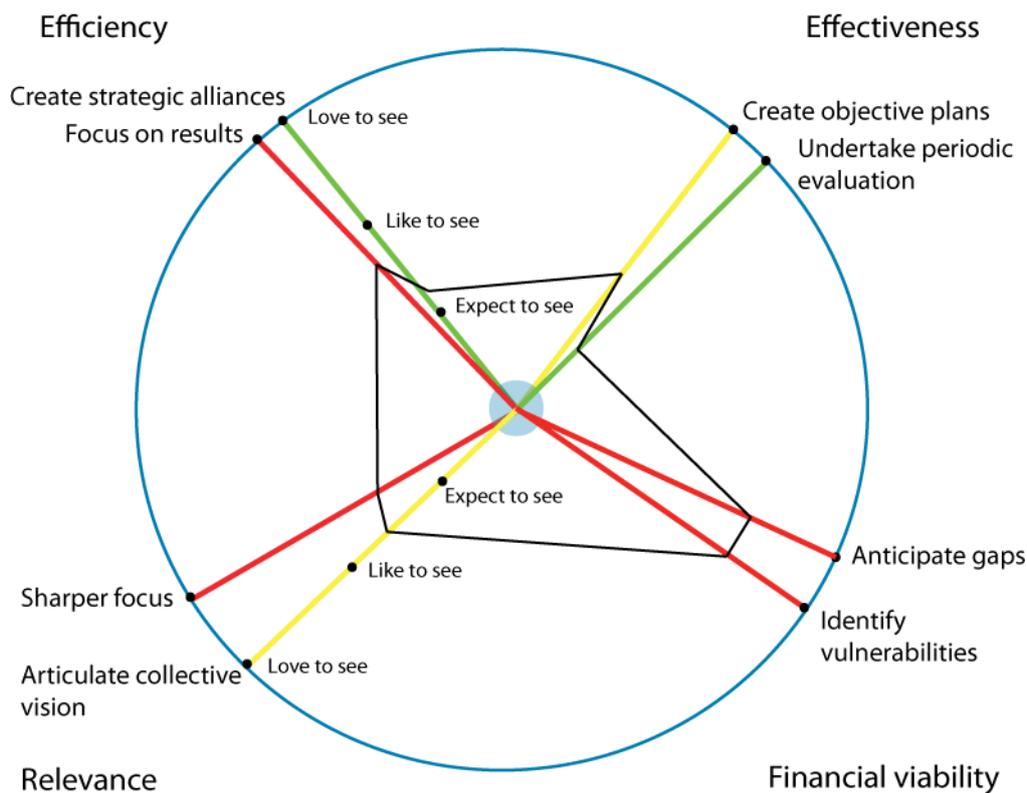
The group then rates the overall level of effort (time, funds, human resources) the organisation would need to dedicate in order to achieve each performance objective, using a colour code reflecting a scale of high (red), moderate (yellow) and low (green). The participatory method used, known as The Socratic Wheel (Chevalier and Buckles, 2012), represents a consolidated picture of objectives and progress markers against which progress can be assessed periodically and finally at the end of 12 months (see Figure 5).

Step 4

The final judgment, and related evidence of achievement of progress markers, is shared with CWY and discussed in detail with CWY Program Managers and staff.

The design combines supplied criteria reflecting the language of effectiveness, efficiency, relevance and financial viability which CWY uses to report to their donor, with specific criteria (objectives) and indicators (progress markers) generated by each partner. This makes it possible for CWY to compile results and report to external donors on changes in the organisational performance of its partners using commonly recognised criteria and standards. At the same time, the process shows flexibility and responsiveness to local knowledge about what organisational performance objectives to pursue, specific priorities, and stakeholder engagement in the evaluation process.

Figure 5 Organisational Performance Wheel



Summary of this example: A private, not-for-profit Honduran development organisation and Canada World Youth (CWY) collaborate on a youth leadership program active in the western part of the country. The NGO has a broader mandate and is responsible for a wide range of development projects, including many in collaboration with smaller local organisations. In June 2011, the NGO developed a tool to monitor progress against objectives it set for itself and to report to CWY on broad categories of organisational performance. The figure shows specific objectives identified by participants, for which graduated progress markers were also developed. For example, with respect to efforts to ensure 'relevance', participants 'expect' to see strategic alliances and a number of inter-institutional agreements by the end of 2011. They 'would like' to see active participation by local groups and other actors in setting organisational priorities, and they 'would love' to see widespread satisfaction with the activities of the organisation. Plans to achieve the objectives were outlined and developed later in greater detail. Lines in the figure (above) joining the various spokes simulate what a completed assessment would look like.

How to Assess Impact?

Answers to questions regarding what needs to be assessed and the purpose that guides the process are fundamental to addressing a central problem in the field of impact studies – the issue of methodological fit. Briefly, our position on this is essentially pluralistic and pragmatic: what matters in the end is that each impact assessment be creatively designed and scaled to fit the situation at hand. And that it combine data collection with multi-stakeholder interpretation and decision-making.

Specialised methodologies developed in a variety of contexts are constantly evolving to address the various dimensions of impact assessment: impacts over time, at a systemic level and through the right causal linkages. These include experimental approaches based on randomised control trials and the use of surveys and questionnaires. Other methodologies draw on Theories of Change, Case Study

approaches and Participatory approaches to impact assessment (Stern et al., 2012). As White notes, “the reports and literature from these different approaches are in general developing in parallel, with rare attempts at dialogue to establish common ground let alone methodological fusion” (White, 2006, p2).

This section makes only general remarks about some major methodological differences among these various options, with a view to presenting a framework developed by Chevalier and Buckles (2012) to bridge some of the gaps. Our goal is to broaden the range of methodological options and ground these in considerations regarding what needs to be assessed, the purpose that guides the exercise, and the people who should be involved.

Different methodological approaches to impact assessment reflect different views on what causality is all about and ways to infer or demonstrate cause-effect relationships. This has resulted in what has been characterised by some as “the causal wars” of impact evaluation (Scriven, 2008; Cohen and Easterly, 2009). In short, battle lines are drawn between quantitative and qualitative approaches and also the suitability of particular methods used to address complex international aid settings and systems.

The origins of the “causal war” can be traced to the popularity in policy circles of the experimental approach to impact assessment most fully developed by Abhijit Banerjee and Ester Duflo (2011). Duflo’s work in this area won her the John Bates Clark Medal, the most prestigious award for economists after the Nobel in economic science. It also attracted the attention of the World Bank and a wide range of quantitatively oriented economists and policy makers, quickly becoming the “gold standard” for mainstream impact assessment. As a result, the experimental approach has starved funding and attention to other equally valid methodological approaches to causation research and evaluation. Critics argue that the level of attention and exclusionary tendencies of the experimental approach in the evaluation field is not merited by the facts or its “special features” (Scriven, 2008).

Characteristically, the experimental approach uses some of the rules for randomised control trials (RCT) common in the medical field to create a **counterfactual scenario** — a description of how things would have turned out differently to the way they actually did. The method requires that part of the target group (a “control group”) be isolated from the effects of the intervention or “variable of interest” (Banerjee and Duflo, 2011; White, 2006). Randomisation is used to give all individuals an equal chance of being in one group or the other. Surveys are then used to collect information on observed changes. This makes it possible, at least in theory, to see the difference between what happened with the program, project or activity, and what would have happened without it, for the same target group.

Critics of the experimental approach question both the practicality and ethics of control groups in real-world settings (Stern et al., 2012; Scriven, 2008; Catley et al., nd). In a complex world, effects are context specific, with no guarantee they will occur in the same way in another future setting or for a neighbouring community or population. This is so true that many fields of intervention (improving rural livelihoods, for example) lack clear and internationally recognised indicators of project performance or other standards against which they can be assessed (Catley et al., nd). Moreover, reliance on large-scale surveys and sophisticated statistical analysis creates real-world gaps between the experimental approach and the resources of many CSOs. It is in part for these reasons that the vast majority of studies conducted using the experimental approach are commissioned by large donor agencies such as the World Bank and led by university-based academics. While CSOs often invest a lot of time and resources undertaking surveys of various kinds, they rarely have opportunities or reasons to work with control groups.

In addition to these flaws, the experimental approach to impact assessment ignores many debates in causal theory (Scriven, 2008), including alternative approaches to counterfactual thinking, a form of evidence-based reasoning that can be found in other fields and professions such as legal practice. In the process of formulating legal judgments regarding responsibility, judges and juries must reflect on

the extent to which the harm observed is justly attributed to the actions of the accused. They go through this fact-finding and reasoning process rigorously, by bringing together evidence and logic to understand how things would have turned out differently *had it not been for the human actions under examination*. Conclusions regarding causation beyond reasonable doubt are reached without relying on random control trials. Similarly, stakeholders in a project or program can determine the extent to which the observed changes would have occurred had the intervention not taken place, keeping in mind the details of the intervention and well-developed evidence regarding the observed changes.

To address this gap, Chevalier and Buckles (2012) developed a method entitled *Attribution and Contribution* to support assessments of the contribution to meaningful change that can be attributed to a specific intervention (action, project, program). The framework helps justify findings and recommendations that follow from a series of considerations – i.e. specification of the intervention, change observed in a domain, the scope of the intervention, the role of other intervening actors and factors, obstacles along the way, how methodical and deliberate the intervention was, what would have happened had the intervention not taken place, and the reliability of evidence provided to answer these critical questions. *Attribution and Contribution* addresses each of these in order, converging around a final judgment on the worth of specific interventions in real settings. The framework, outlined and briefly illustrated below (Box 6), offers a reasoned response to “the attribution problem” and to broader debates concerning the concept of causation in the social sciences. It also shows how to bridge the gap in conventional methods between data collection and interpretation and decision-making needed to make sense of complex settings, and go beyond the limitations of mainstream evaluative methods (involving surveys, descriptive statistics, interviews, focus groups and storytelling).

Box 6: HIV/AIDS prevention and treatment (using *Attribution and Contribution*)

Step 1: Intervention, objectives and partners

*Attribution and Contribution begins with a description of the **intervention** (action, project, program), its main **objectives** (fixed or adjusted over time), the corresponding **time frame** and the implementing **partners**.*

The Unitererra IVC program managed by CECI and WUSC in Burkina Faso engages Canadian, national and South-North volunteers in work with partner organisations to reduce the incidence of HIV/AIDS and sexually transmitted diseases. During a two-year period (2009 to 2011), Unitererra brought 62 Canadian volunteers (38 women, 24 men) to build the capacities of seven different partner organisations in areas ranging from HIV/AIDS prevention (information, education, behaviour change, communication, counselling for voluntary HIV testing) to gender equity awareness and the provision of support and care services (palliative care, helping orphans and vulnerable children achieve financial autonomy, etc.). The Unitererra program and its volunteers have also sought to strengthen organisational governance and efficiency by bringing improvements to strategic planning, the management of human resources, accounting for resources, mobilising resources and managing knowledge (through better secretarial work, archival methods, data base management, communications, monitoring and evaluation).

Step 2: Observed change

*Participants in the assessment discuss the relevant **observed changes** in the domain. Based on the available evidence, they form and justify a judgment regarding the magnitude of the observed changes –*

whether there has been major, moderate, small or no progress, or whether the situation has worsened.

Based on studies using appropriate methods including surveys, interviews and self-reporting, the participants in the Unitererra program noted that since 2009, various changes have been observed.

Increased outreach

Of the three partners working on prevention, all have increased outreach, each in their own way. Some offered counselling for voluntary HIV testing. Another created a theatre forum involving 12 people living with HIV (PLHIV) who gave performances for a total audience of 12,192. A third built a more efficient electronic management and provisioning system for pharmaceutical products used by PLHIV.

Enhanced support and care

Using a novel approach that incorporates ergotherapy and psychosocial support, three of the six partners offered support and care services to 6,217 PLHIV. Palliative care services reached 70 women and 30 men living with HIV, and another 228 received nutrition counselling. As for measures to promote financial independence, reports indicate that the theatre forum raised USD 2,000. Another 31 PLHIVs secured micro-loans totalling USD 6,400, 80% of which were reimbursed within the expected period. Income generating activities helped 13 orphans and vulnerable children achieve greater financial autonomy. Four of these children have seen their earnings from a soap cooperative jump from USD 1,000 to USD 2,000 in 2011.

Improved leadership and governance

All partners mentioned improvements in leadership and governance. For some, this meant producing high-quality financial reports and delivering them on time to decision-makers; for another, a review of its by-laws helped to clarify the roles of directors and staff; for a third, producing a news bulletin four times a year increased its visibility and accountability to supporters. All partners indicated that knowing how to better mobilise financial resources and diversify sources of earnings contributed to greater organisational coherence and project sustainability.

Greater gender equity

Two of the partner organisations increased substantially the number of women in key decision-making positions within their organisations. Another delivered gender equality workshops to 30 elected representatives and 50 association leaders. More broadly, partner awareness of gender equity issues increased.

Enhanced engagement in policy dialogue

Partners reported being better prepared to participate in political dialogue, with a view to influencing policy-making and the development of important HIV/AIDS services such as palliative care.

At a broader level, the seroprevalence rate in Burkina Faso has gone down from 1.6% in 2007 to 1.2% in 2010, according to official UNAIDS statistics. Sexual behaviour seems to be changing as well. In 2010, about 70% of adults aged between 15 and 49 said they used a condom the last time they had sexual intercourse. These reflect positive changes for the country as a whole.

Step 3: Intervention scope

*Further specification of the intervention scope is needed to begin to build understanding of the link to observed changes. Was the intervention the sole contributor to the observed change, or did it play a **role** jointly or in parallel with other interventions or intervening factors? Did the intervention contribute directly to the observed change, through **nearness** of cause and effect, or did it act indirectly or somewhat remotely, several steps removed from the final effects? What was the **scale** of the intervention — large, moderate or small? How important were the **obstacles** blocking progress? Were they major, moderate or minor, if any? Responses to these questions can be consolidated in a figure using a rating process, for reference later in the assessment process.*

Discussion among partners of the Unitererra program concluded that it had made an indirect contribution to the observed changes by jointly interacting with others. The scale of their intervention was moderate and moderate obstacles were encountered along the way (delays and

Discussions by Unitererra stakeholders concluded that the program and its partners were methodical in bringing about the observed results, and that the evidence to show this is strong and verifiable. Program planning documents created through the RBM process show the rationale for the interventions and logical steps to be taken. They

interruptions in funding that hampered the implementation of project activities as planned).

Step 4 and Step 5: Default scenario and overall results

Determine the extent to which the observed change would have occurred anyway had the intervention not taken place, keeping in mind answers to previous questions. This is the default scenario, or counterfactual statement. Indicate whether there would have been major, moderate, small or no progress, or whether the situation would have worsened. Also, discuss and rate the overall result of the intervention. This is the difference between the observed change and the default scenario — the distance between the two prior ratings on the vertical line in the centre of the diagram (see example in Figure 6). Is the overall result positive or negative? Is it significant, modest or limited?

Assessment of the default scenario by the partners of the Unitererra program concluded that without the help of Unitererra volunteers, partners would not have had the capacities and staff to offer new services such as palliative care, the theatre forum and counselling for voluntary HIV testing. The visibility of existing networks would have remained low were it not for Unitererra's assistance in strategic planning, network resource mobilisation and institutional reform. Financial viability would also have been compromised for three of their partners were it not for support to income generating activities provided by volunteers. Overall, organisations are now more competent in their own priority areas and better equipped to mobilise the financial resources they need to implement and sustain project activities. Given the distance between the default scenario (involving small or no progress) and the observed changes (moderate changes), the overall results are positive and significant.

Step 6: Methodology of the intervention

Another key consideration in legal reasoning regarding responsibility for a situation concerns the methodical nature of the intervention (expressed in the concept of rational premeditation). When applied in an evaluation context, four questions help to determine this:

1. *To what extent did the intervention use **effective** methods, i.e. rational steps and credible means, to achieve the observed results?*
2. *To what extent did it make an **efficient** use of available resources (human and material)?*
3. *Did the intervention achieve results through steps and adjustments that were **deliberate**, or were results partly or fully **accidental**, i.e. obtained through unintended actions?*
4. *How **verifiable** is the evidence used to answer all preceding questions? Is it generally sound, incomplete, or rather weak?*

provide detailed information on M&E indicators and ongoing performance measurements using the Logical Framework and reflect careful and deliberate thinking about cause-effect relationships. Evaluation data generated through iterative reporting by volunteers, organisations and a sectorial committee also support the conclusion that

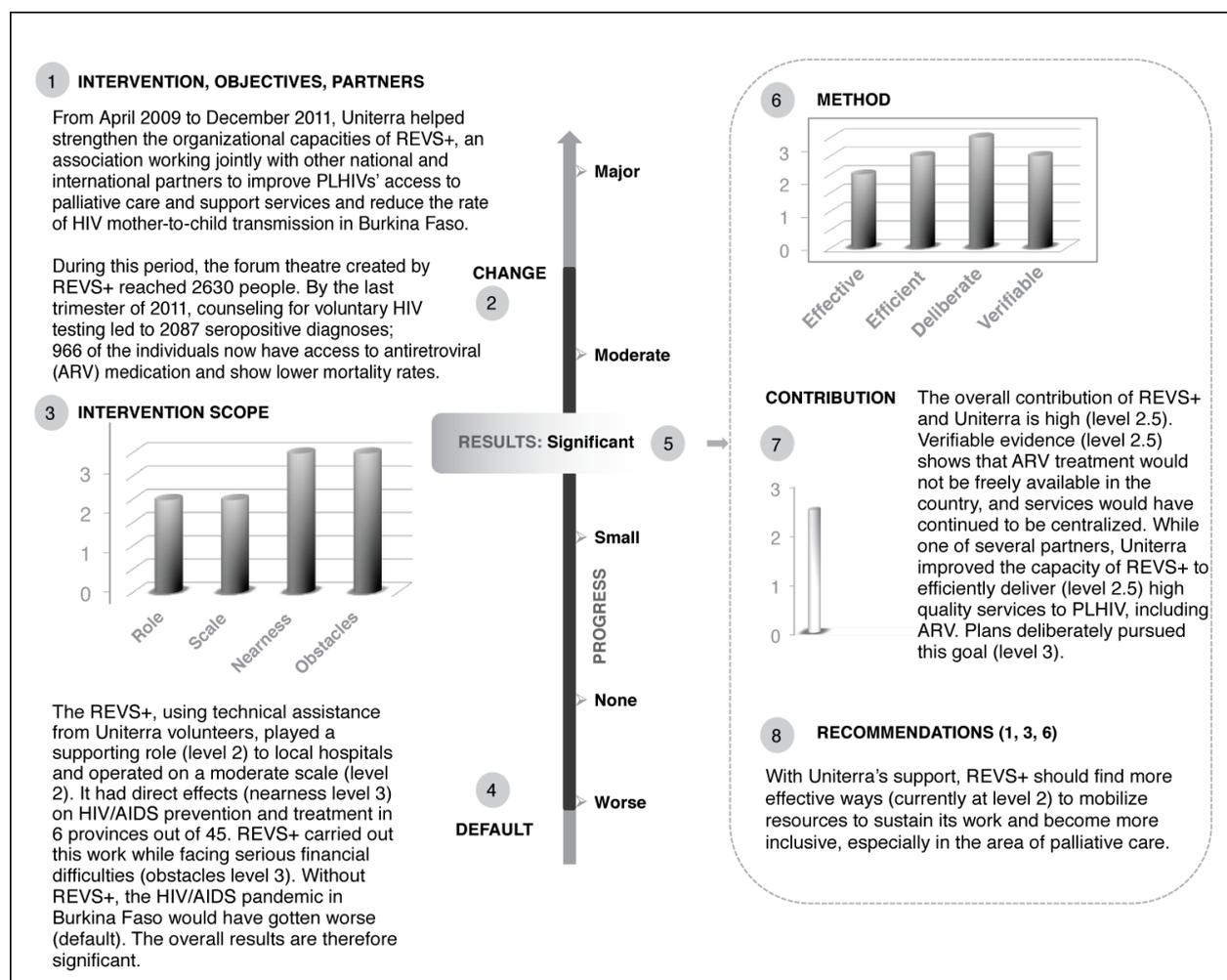
the methods of the intervention were effective and deliberate. On the matter of efficiency, however, things could be improved in two ways: by adding more national volunteer placements to the program and by clearly identifying the people responsible for planning, M&E and reporting activities. Figure 6 consolidates the results of these deliberations and synthetic ratings on each question.

Step 7 and Step 8: Contribution and recommendations

Assessment of the overall contribution that an intervention has made to observed changes makes use of the answers to all previous questions consolidated in figures and associated ratings (Figure 6). It also lays the groundwork for discussion of the implications or recommendations that follow from the assessment.

When reaching these last steps, the HIV/AIDS sector committees in Burkina Faso concluded that Uniterra made a significant joint contribution to the introduction of important new HIV/AIDS services and the development of organisational capacities in the sector. Three recommendations followed. First, some adjustments to project priorities are in order, to more directly reinforce Burkina Faso's work in the area of HIV/AIDS prevention. Second, the results of project activities such as voluntary HIV testing and the prevention of mother-to-child HIV transmission (PMTCT) should be consolidated, preferably through contributions from national volunteer placements. Third, given the current decline in funding for HIV/AIDS projects, Uniterra should continue to strengthen its contribution to capacity building in the area of resource mobilisation and the development of strategic North-South partnerships.

Figure 6 Uniterra impact assessment (using Attribution and Contribution)



Conclusion

This paper has argued that there can be no all-purpose impact assessment methodology that addresses all possible situations where efforts to build understanding about the impacts of IVC programs are needed. The goals and activities people want to monitor and evaluate are as varied as the projects and programs in which they are involved. So, too, are the baseline conditions against which the activities are assessed. Navigating through the methodological options per se is useful and necessary only when some work has already been done to provide answers to questions regarding what needs to be assessed, for what and whose purpose, and with whom. These questions can be answered sequentially, but are more likely to need tentative answers that are adjusted in light of answers to later questions along the way.

The design of impact assessments needs to match the purpose, the level of complexity and the constraints (time, resources, skills) found in each situation. To do so, it must be flexible, meaning that it:

- can **mix** tools and **adapt** them to **match** the situation;
- mesh and integrate both **qualitative** and **quantitative** thinking and findings;
- is **practical** and **time-efficient** (avoiding exhaustive data and text-heavy reports);
- can be **scaled up or down**, to meet needs and existing constraints (financial and human resources);
- can generate both project and **higher level findings**.

Our argument is more than a call for mixed methods, however. While many methods have something to contribute to developing appropriate impact assessment designs for IVC programs, including the hard experimental approach, in our view, further innovation is needed to ensure that impact assessments:

- bring together the processes of fact-finding and the construction of meaning in complex settings, aspects of evaluative thinking currently separated in mainstream methods (surveys, descriptive statistics, interviews, focus groups and storytelling);
- serve purposes other than upward accounting for resources and results, to include learning and **adjusting plans** for ongoing or future actions, or sharing a project or program story to **inspire** others with lessons learned;
- support an **ongoing feedback**, action-reaction loop (as in medical practice), to acknowledge learning from failure and constantly address the “So what?” and the “Now what?” questions;
- factor in the effects of uncertainty and **complexity**, including **multiple stakeholder contributions** to observed results;
- encourage participatory approaches that facilitate **interactive engagement** and mutual learning and accountability. This is a principle often ignored in self-evaluations (isolated from challenges by others) as well as in evaluations by proxy (assessing one’s work by reporting on the results of one’s partners) or carried out by third party experts (using survey, interview and narrative data).

These additional principles provide direct inspiration to the design of the impact assessments briefly presented in this paper and the related tools developed by the authors (Chevalier and Buckles, 2012 and www.participatoryactionresearch.net). They suggest that collaborative gathering and interpretation of evidence, sense-making and decision-making must be built into the methods used to assess impacts. This is key to developing a more meaningful understanding of the difference made by IVC programs.

References

- Banerjee, Abhijit and Ester Duflo (2011) *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*, Public Affairs Books, Philadelphia
- Catley, Andrew, John Burns, Dawit Abebe and Omeno Suji (nd) *Participatory Impact Assessment: A Guide for Practitioners*, Feinstein International Centre, Tufts University, Medford
- Chevalier, Jacques M. and Daniel J. Buckles (2012) *Participatory Action Research: Theory and Methods for Engaged Inquiry*, Routledge, UK
- Cohen, J. and W. Easterly (2009) *What Works in Development? Thinking Big and Thinking Small*, The Brookings Institution, Washington D.C.
- Christie, J. (2008) Evaluation of Development Action – Why? of What? for Whom? The Challenges of Evaluation, as seen by some Canadian CSOs. Report for the Canadian Council for International Co-operation
- Davies, Rick (2005) 'The "attribution problem" problem', Blog entry, December 16 (<http://richardjdavies.wordpress.com/>, accessed July 4, 2012)
- Earl, S., F. Carden, T. Smutylo (2001) *Outcome Mapping: Building Learning and Reflection into Development Programs*. International Development Research Centre, Ottawa
- Easterly, William (2006) *The White Man's Burden: Why the West's Efforts to Aid the Rest Have Done So Much Ill and So Little Good*, The Penguin Press, London
- Jackson, E.T. and Associated (2007) *Mid-Term Evaluation of the Volunteer Cooperation Program*. Report for the Canadian International Development Agency
- McBride, A. M., B.J. Lough, and M.S. Sherraden (2010) Perceived impacts of international service on volunteers: Interim results from a quasi-experimental study (CSD Research Report 10-22). Washington University, Center for Social Development, St. Louis
- McLaren, K. and P. Turcot (2006) *Canada World Youth Impact Assessment*. Synthesis Report for Canada World Youth
- Mwathi Mati, Jacob and Helene Perold (2012) 'Youth Volunteer Exchange Programmes in Southern and Eastern Africa: Models and Effects', Unpublished report by Volunteer and Service Enquiry Southern Africa (VOSESA), Canada World Youth and Southern Africa Trust
- Owens, Joanne (2011) 'Measuring System Dynamics in Cuso International Programming', Cuso International, unpublished report
- Scriven, Michael (2008) 'A Summative Evaluation of RCT Methodology: An Alternative Approach to Causal Research', *Journal of MultiDisciplinary Evaluation*, vol 5, no 9, pp11-24
- Sherraden, M., B. Lough, and A. M. McBride (2008) 'Effects of international volunteering and service: individual and institutional predictors', *Voluntas*, vol 19, no 4, pp395-421
- Stern, E., N. Stame, J. Mayne, K. Forss, R. Davies, B. Befani (2012) 'Broadening the Range of Designs and Methods for Impact Evaluations', Department of International Development (DIFID), Working Paper 38
- Tiessen, Rebecca and Barbara Heron (2012) 'Volunteering in the developing world: the perceived impacts of Canadian youth', *Development in Practice*, vol 22, no 1, pp44-56

Watson, C. (2008) Literature Review of Impact Measurement in the Humanitarian Sector, Feinstein International Center, Medford

White, H. (2006) Impact Evaluation: The Experience of the Independent Evaluation Group of The World Bank. Independent Evaluation Group, Washington, D.C.