

Defining and refining effectiveness: applying narrative and dialogue methods in aid monitoring and evaluation

Juliet Willetts, Helen Cheney and Paul Crawford

Abstract

In this chapter we argue that definitions of ‘effectiveness’ should be negotiated and that this can be assisted through the careful selection of monitoring and evaluation methods. A flexible view of ‘effectiveness’ is necessary because the world is complex and ever-changing, and even more importantly because the concept is contended. Inevitably, different stakeholders in development aid hold different perspectives about what ‘effectiveness’ means in their context.

We propose that the use of certain types of qualitative methods in monitoring and evaluation is an important way to promote dialogue on the different definitions of effectiveness between stakeholders. In support of our argument, we demonstrate how two open-ended inquiry methods were applied in the field to first define and then refine the meaning of effectiveness. We also apply a validation framework to test the quality of these methods and to generate insight into the strengths and weaknesses of their application in the field. A key strength of these methods is found to be their attention to capturing diversity of perspective. In addition, both the narrative and dialogue methods are seen to indeed facilitate the negotiation of meaning of effectiveness between the different project stakeholders. We conclude that a thoughtful and flexible approach to monitoring and evaluation that incorporates such qualitative methods enables effectiveness to be defined and refined, and is conducive to more appropriate, better managed aid.

Introduction

Effectiveness as a fluid, contended, multi-perspective issue

‘Effectiveness’ is a challenging concept in the development sector. This is partly because it is fluid and evolving, where meanings may be developed through interaction and learning in a project or program (Kaplan 1999; Smutylo 2001). It is also challenging because of the different perspectives, values and interests held by stakeholders. Aid effectiveness relates closely to aid objectives, which are often contested. There is not, and will never be, universal agreement on the relative importance of different aid objectives. Therefore the evaluation of aid, which involves investigation of aid ‘effectiveness’ as it relates to those objectives, also attracts such disagreements.

In this chapter, our focus is on assessing effectiveness of *projects* rather than effectiveness of programs and organisations. The qualitative methods we discuss could however also be used for these other kinds of evaluations. Aid effectiveness in the project context relates to the extent to which the project’s objectives were achieved, and the extent to which the project’s objectives were the ‘right thing’ - appropriate for the context and the best or the most strategic intervention that could have been done (Checkland 2001; Crawford 2004).

Aid project stakeholders include, as a minimum, the project beneficiaries, the project implementers (project staff, as well as local, regional and head office staff, and perhaps including an office in the donor country) and the project donor. Different perspectives will inevitably exist between and within these groups about what is needed at the stages of project planning, implementation and in evaluating project effectiveness.

A concern for this plurality of perspectives raises the question of whose perspective on 'effectiveness' matters. This can result in a strong focus on the beneficiary perspective (Barry & Fourie 2001; Chambers 1997; Fowler 1996). While this focus is valid and indeed essential, there are other perspectives that also matter. Therefore what is also needed are explicit ways to negotiate between different stakeholder perspectives.

This negotiation requires the definition of a project's effectiveness to be an evolving process. That is, the interpretation of 'effectiveness' needs to be iteratively defined and refined within the project and the evaluation cycles. The stakeholders, as learners in this process, can be expected to evolve and change their perspectives over time. This evolution takes place as the result of dialogue and repeated 'accommodations' made within and between the different stakeholder groups. Agreement on an appropriate definition or indicator may be reached by a set of stakeholders and be acted upon within a particular project cycle. However this agreement is best seen as contingent and the definition of effectiveness should be expected to change at some future point. Learning may take place, circumstances may change and new negotiations among stakeholders will likely be required. Defining and refining effectiveness in this way makes sense in a project context where a long-term on-going relationship is sought between stakeholders and where there are concerted efforts to overcome power asymmetries. This appears to be a growing trend, with shorter term engagement acknowledged as a less successful means of aid delivery (Kelly & Chapman 2003; Madeley 1991).

Defining and redefining effectiveness during monitoring and evaluation processes

Monitoring and evaluation may be used for different reasons. Monitoring and evaluation often includes processes to promote reflection on the activities and effects of a project. This enables learning for the various stakeholders and potentially facilitates new or more appropriate interpretations of effectiveness. Accountability is also a key aspect of aid monitoring and evaluation. A narrow definition of effectiveness may lead to an evaluation that is only meaningful to stakeholders holding that particular definition, with less potential for accountability to stakeholders holding alternative definitions. For instance, in cases where the donor's definition of effectiveness requires quantification of particular outcomes or impacts, accountability to the project beneficiaries and project implementers is limited if these groups align with different definitions of effectiveness. Moreover, this approach may not recognise or capture unintended effects. If, however, effectiveness is defined more broadly and with flexibility, then broader accountability is possible.

Research methods used in monitoring and evaluation include both quantitative and qualitative approaches, with mixed method approaches becoming increasingly popular.

Here we argue that the thoughtful inclusion of qualitative approaches in monitoring and evaluation is important where the definition of effectiveness is to be negotiated.

The value of qualitative approaches (Spencer et al. 2003) is that they:

- are concerned with meanings, especially the subjective meanings of participants;
- are sensitive to the social context of the study;
- retain complexity and diversity in the analysis; and
- attend to emergent categories and theories rather than reliance on *a priori* concepts and ideas.

Qualitative approaches are able to capture diverse perspectives as well as the links and patterns between them. They are also able to enhance communication among stakeholders in a monitoring and evaluation process. Participatory and open-ended methods are especially good for supporting diverse perspectives and pay attention to unintended (both positive and negative) changes.

Narrative and dialogue methods

The *narrative* and *dialogue* methods described in this chapter are participatory forms of qualitative inquiry. *Narrative* methods involve the use of stories, in data collection and as a mode of data analysis. Narrative approaches to data collection encourage participants or groups to give long “story-like answers to a question or questions or describe their thoughts or memories around or in response to a topic” (Wadsworth, 1984, p44). Their value in the development context is that they allow participants to utilise their own meaning making systems within their own language, culture and prevailing narrative techniques (Reissman 1993). This is important because there are likely to be different languages and cultures between donors, implementers and recipients. In this chapter we use a narrative method to describe the collection and analysis of stories told by individuals. These stories may belong to the group but have been told in the context of an individual interview.

Dialogue approaches vary widely, for example from how they are used in anthropology, to conversation analysis to group interviews. Here we use the term to denote a process of group meaning making (as compared with individual stories). We use dialogue methods to bring out the interpretations of groups through “exchange of experiences or ideas or views” (Wadsworth, 1984, p44). Within a group, participants give opinions, tell stories and respond to one another. They may add to or contradict the views of others. A dialogue method pays attention to differences in perspective rather than looking for consensus or a group response.

Both of these approaches are sensitive to the voices of participants and the diversity of views in a particular context. They are therefore suited to negotiating the meaning of effectiveness in the context of a project evaluation. Two case studies are provided below to demonstrate these methods in practice.

Case study analytical approach

This section discusses how we have chosen to approach the two case studies presented in this chapter. Firstly, following clarification of the case study context and setting, we have provided significant detail on the evaluation method design. This is both to alert the reader to how small details in the approach may have a significant impact on the results collected and to facilitate replication of such methods by others.

We then apply a quality framework to the case studies. This is to provide a deeper level of reflection and insight to the presentation of the case studies, and to demonstrate key aspects of what quality means for qualitative methods. The quality framework we have chosen (Spencer et al. 2003) was developed through the synthesis of a wide variety of different quality frameworks. It is called “Quality of qualitative evaluation: A framework for assessing research evidence” (Spencer et al. 2003). This quality framework is comprehensive and widely applicable to different qualitative approaches. It contains four guiding principles and eighteen assessment questions which may be applied retrospectively. The four principles are that the qualitative evaluation should be contributory, defensible in design, rigorous in conduct, and credible in claim. The assessment questions (in which these principles are embedded) concern all stages of the evaluation research process (sampling, data collection, analysis, reporting). In addition the assessment questions include three other important aspects, reflexivity and neutrality, ethics and auditability. We conclude each case study with a discussion of how the methods assisted negotiating definitions of effectiveness.

Case study 1: The use of a narrative method to define and refine project effectiveness – trialing the ‘Most Significant Change’ Technique in Laos

Context and setting

A pilot project to trial the ‘Most Significant Change’ (MSC) monitoring technique resulted from the observation by senior management of Adventist Development Relief Agency (ADRA) Laos of a growing trend, especially among their traditional donors, of questioning the effectiveness of non-governmental organisation (NGO) implemented projects. In response, ADRA Laos initiated more systematic and participatory monitoring and evaluation of its development program. This was needed not only to improve the quality and performance of their development activities through provision of relevant management information, but also to improve their ability to demonstrate the organisation’s program effectiveness to key stakeholders. In 2003-4 they decided to pilot the MSC qualitative monitoring system for nine months to determine its ability to address these needs.

The MSC pilot project was implemented for three of ADRA Laos’ projects: two that work in tandem in the Luangnamtha region in the north of Laos and one in the southern region of Attapeu. The projects were focused on water, sanitation and health and hygiene education.

At the outset of designing their MSC system, staff defined five goals:

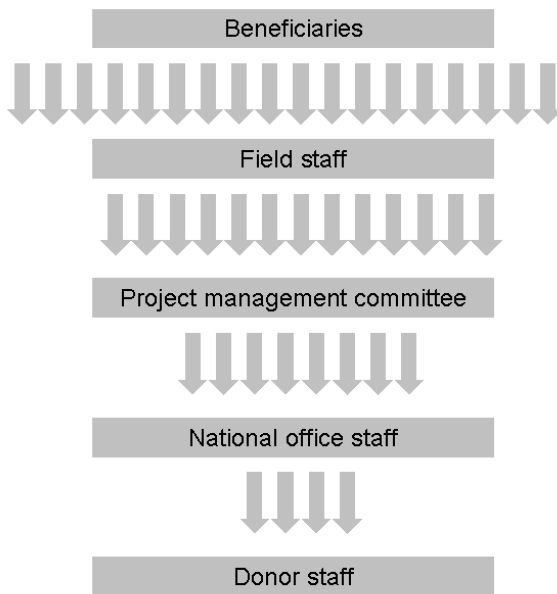
- To increase the participation of stakeholders in the monitoring and evaluation process.
- To develop the analytical skills of field staff.
- To improve ADRA Laos’ ability to determine overall impact of projects.
- To improve ADRA Laos’ understanding of how projects are interacting with beneficiaries.
- To improve ADRA Laos’ project management.

One of the authors of the chapter, Juliet Willetts, conducted the external evaluation of this pilot study at the end of the nine-month period. It is from this meta-evaluation study (Willetts, 2004) that much of the information presented below is drawn.

Monitoring method design

The design of the MSC system for ADRA Laos was based on an investigation of MSC in other contexts. Figure 1 shows how stories were collected and selected by different stakeholders. Local field staff collected one or two stories in any villages that they were working in during the month. Four domains of change delineated types of change. These domains were changes in health, changes in behaviour, negative change and any other change. These stories of change were collected in interviews conducted by field staff with beneficiaries and from field staff observation. In interviews, the questions about change were: “What is the best change since ADRA’s project?” and “What is the biggest problem due to ADRA’s project?” However, these specific questions were used to varying extents, as is discussed later in our assessment using the quality framework.

A specific instrument was designed to document the MSC stories and included both the details of the change and why it was significant to the villager and the staff member. Each month, the project management committee (PMC) for each project (several of the field staff and the project director), read and analysed the stories. Following a process of



lengthly discussion to reach consensus, the PMC selected one story reported by beneficiaries and a story based on their own observations of a community in each domain (a maximum of eight stories). These selected stories were then sent to the main office of ADRA Laos in Vientiane, including documented reasons for their choice. The PMC also included recommendations for actions by field staff in response to stories and issues they raised. A committee at the main office in Vientiane met monthly to deliberate on the stories, to select one story in each domain, and send feedback to the PMC, commenting on their recommendations. The final stories selected from six months of monitoring were sent to ADRA Australia.

Figure 1: Story collection and selection process

Assessment using the quality framework

Key strengths of the field application

The eighteen assessment questions of the quality framework were applied to the MSC pilot project in Laos. Through this process five main strengths were revealed.

- The preservation of original data to maintain a clear evidence base.
- The immediate new insight that the data (stories) provided to staff.
- The capture of the basis of evaluative appraisal in the technique, that is, ‘why’ a story was considered significant.
- The well-organised data collection and documentation protocol.
- The strong diversity of views captured, particularly the inclusion of a negative ‘domain’ to capture criticisms.

These strengths contributed greatly towards the ability of the method to evolve meanings of effectiveness by the project stakeholders. Each is described in more detail below.

The MSC method builds credibility of the findings by explicit preservation of the original stories in an intact form. This is in line with current thinking on narrative methodology:

“[p]recisely because they are essential meaning-making structures, narratives must be preserved, not fractured, by investigators, who must respect respondents’ ways of constructing meaning and analyze how it is accomplished.” (Riessman 1993: p4).

The second strength identified through the assessment using the quality framework was that the stories provided immediate new insight to staff, that is, the monitoring system served to contribute to new knowledge. In fact the effect of the stories on both management and field staff was transformative. One management staff reported that use of the method “has created deep changes in people’s thinking amongst the staff...[and]... has forced in-depth development thinking” (Willets 2004, p6). Field staff who were involved in collecting stories reported that through using MSC “we can measure if work done is fruitful or not, did they understand, did they practice or not?” (Willets 2004, p6). It is these new insights that permitted new perspectives on effectiveness to evolve.

The quality framework asks “how clear is the basis of evaluative appraisal?” in terms of what judgements have been made, their basis and by whom. In the use of MSC in Laos, this aspect of the method was clearly documented. The analysis involved either individuals or groups determining which story (from a selection of stories) is the most significant, and articulating ‘why’ they believe that story to be most significant. Field staff were the first to consider the original set of stories and subsequently sent a smaller number of stories to management staff, who in turn chose, and sent a smaller selection to the donor organisation. At times complex debate about the selection of stories was not well-captured and documented. However the vital part of the MSC process for the purposes of defining and refining effectiveness, was the debate itself and new insights it may have triggered for those involved. The real importance of such debate lies in how

individual values and assumptions are revealed. The debate brings them to the surface so that each person's subjective judgement and mode of reasoning is made transparent.

According to the quality framework, data collection needs to be systematic, well-documented and described, as well as making sure that depth and richness are achieved. In Laos, the data collection was well-organised and the data collection instrument ensured that many details of the context (such the person, age, place, story collector, date etc.) were recorded. In addition, when stories were appraised by different groups, the reasons for their choices were added below the original story. The actual process of eliciting and capturing stories is discussed in more detail in the next section on limitations, as while overall this was done well, it also presents challenges in the use of the method.

The use of MSC was excellent in its ability to explore diversity of perspective and content, another area of importance in the quality framework. Diversity was achieved through collection of stories from a wide variety of sources, including both beneficiaries and observations made by field staff. The four domains of change included negative change and stories collected in this domain were especially valuable. For example, stories about situations where the project intended to catalyse a change (for instance, boiling water before drinking) but no change had taken place, and stories about problems or challenges faced within the project or its effects in the community. This explicit attention to negative change was extremely important in provoking reflection in both field and management staff about their roles, the purpose of their work, and the perspectives they held on effectiveness.

Limitations of the field application

Application of the quality framework revealed four areas of concern in the use of this narrative method. While previously identified by Willetts and Crawford (2007) the use of the framework has enabled explication of these limitations. These four areas were the lack of a systematic mechanism to respond and follow-up issues learnt through the monitoring system, haphazard sampling, limited research skills of some field staff in collecting stories and finally, ethical issues with regard to preserving anonymity and obtaining consent.

The monitoring system was new and this may explain why its design did not include explicit mechanisms for feedback and response, and why therefore this area received inadequate attention. Both field staff and management staff reported that only some stories were followed up and others contained issues that were left unattended. It became clear that there was a need to integrate the MSC monitoring system with other monitoring systems to ensure follow-up. Also, the outcome of the selection process at the management and donor level was at times not well-communicated back to field staff and not at all to the beneficiaries who provided the stories. Both these issues of follow-up and feedback could be easily remedied.

The selection of informants to provide stories was a flaw in the sampling process as this needed to be more systematic. Certain groups of people may have been omitted from the

story collection process and this weighed against the practical benefits of collecting stories in an ad hoc, opportunistic basis as had been done. The main group omitted was young adult workers who are mostly in the rice fields, and were therefore often not present in the village at the time of story collection. Further planning and thought would determine the best way of ensuring that all groups are represented, or at least that staff are given the responsibility to track the demographic characteristics of villagers interviewed and which villages have been covered each month.

The quality framework includes an assessment question focused on reflexivity. The effect of their own role in eliciting stories was appreciated by some field staff, but not others, despite specific training in this area. For instance some field staff were unaware of the problems created by asking leading questions rather than letting the villager direct the conversation and choose their story of significant change. The field staff story collection task is one that requires well-developed interpretative research skills to elicit stories that have integrity. This is difficult to manage in the development setting where staff have limited formal research training. On-going training and mentoring in the use of such qualitative methods is required.

Two important ethical issues arose in the use of this narrative method in Laos. The first concern was that anonymity of the informants was not well-managed. As mentioned earlier in regard to data collection, all stories were annotated with who told the story. This could have potential ramifications for an individual if for example a story was used in a way that damaged them or the community, or, if personal information was expressed that was considered private. What makes this issue even more problematic is a second ethical issue that often informants were not clearly informed when a story was being collected. For instance the story collection might occur as a casual conversation (staff usually wrote down the story after the event) and in this situation informants may not have even been aware that personal information was being collected from them. Promotion of open dialogue to elicit meaningful stories (particularly negative ones) could be better and more ethically achieved through trust and openness rather than covert collection of information. It is very important that evaluation processes are able to balance the need for accountability with the need for learning so that information is not withheld in fear of punitive consequences. Some discretion about reporting the “bad” news is necessary in contexts where there are clearly asymmetrical power relations in operation.

Explanation of how effectiveness was defined and refined

In relation to the MSC monitoring system, field staff, management staff at the country level, and donor staff in Australia all described changes in how they perceived the projects, their role and their conception of effectiveness. These changes have already been touched upon and are discussed more explicitly below.

Field staff’s view of the aims of the water and sanitation project shifted through use of the monitoring method. One staff reported how in the past they had considered that if they had arrived, dug a well and showed the community members how to operate and maintain it, that their job was done. Use of the monitoring system however, had made it

clear to this person that this was simply not enough. Importantly, it became evident that various risks and assumptions that had not been considered at the time of project conception were playing a significant role in determining the extent of project success. For example, one story collected through MSC revealed that the taste of the water was found to be unpleasant in one community, therefore the villagers had returned to drinking the contaminated surface water. In response to such stories the field staff definition of effectiveness evolved from a narrow focus on delivery of health education and provision of water infrastructure, to a broader focus on genuine and lasting changes in behaviour achieved through strong and trusting relations between project staff and beneficiaries.

Management staff shifted their perspective on effectiveness towards one that was more grounded in the needs and values of the project beneficiaries. Actively debating which stories represented the most 'significant' change played a valuable role in challenging staff's individual values and assumptions about what constituted effective development aid and what they perceived their organisation's purpose and role to be. The shift saw a move away from a 'delivery' model of aid to one more focused on relationship building and more two-way exchange between the aid organisation and the communities they served.

At the donor level, staff indicated that MSC helped create dialogue about their purpose and the value of being exposed directly to details of what happens at the village level, which was often not a part of many of the staff's experience. For instance the appraisal and review panel (a panel external to ADRA Australia who review new projects) indicated that the sharing of the MSC stories, particularly those that described unintended change and lessons learnt, was extremely valuable in helping them to comprehend the reality of aid delivery in the field, and to guide ADRA's investments appropriately.

Case study 2: The use of dialogue to define and refine the concept of 'well-being' in Ecuador and Peru

Context and setting

The final evaluation of a jointly funded World Vision/AusAID Andean Permaculture Project (2001-2006) was conducted by one of the authors of this chapter, Juliet Willetts. This Permaculture Project was implemented by World Vision Ecuador in communities in three different regions (Achupallas, Pilahuin and Saquisili), and by World Vision Peru in eleven communities in the Lamay region. The Permaculture Project had two main project objectives: to improve food security and to improve well-being.

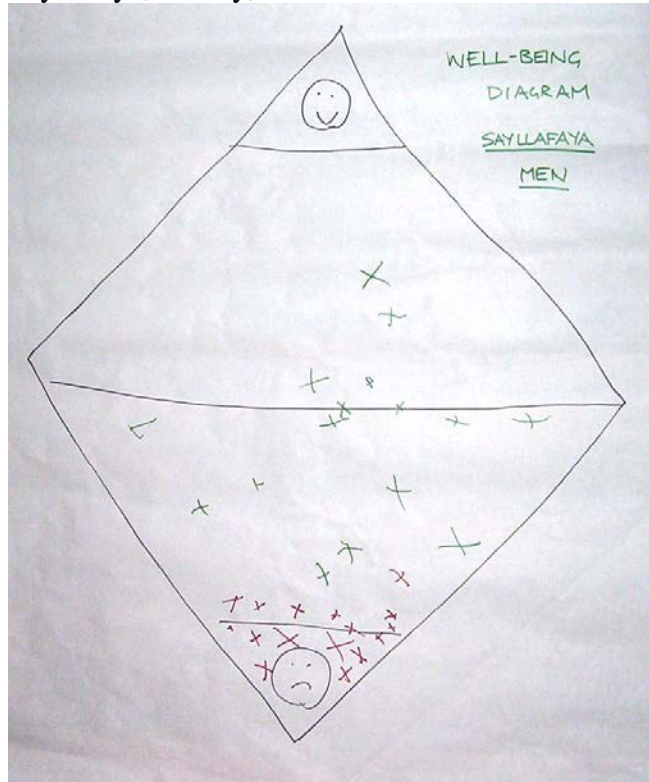
The evaluation covered the efficiency, effectiveness, sustainability and appropriateness of the project. In relation to effectiveness, the evaluator explicitly sought the beneficiary perspective on well-being, rather than adopting the donor or implementer definition of well-being. A separate, but related, method was used to assess changes in food security (not covered in this chapter).

Evaluation method design

A dialogue method that uses a visual aid to guide group discussions was used to examine the issue of beneficiary well-being. This method was developed by others as part of participatory action learning approaches in Uganda and India (Mayoux et al. 2005). Such approaches are useful to promote active participation of the poor and those with little or no formal education (Mayoux et al. 2005). The application of the method in this evaluation involved some quantification of the responses however we focus on the qualitative aspects of the method in this chapter as these are most relevant to defining and refining effectiveness.

The group discussions were conducted with both women and men separately in five communities (three in Ecuador and two in Peru) and the groups varied in size from three to thirty. A sequence of steps was followed, though the order was varied to accommodate spontaneous and informal environment and to encourage participants to lead the discussion.

Figure 2: Well-being diagram for men in Sayllafaya, Lamay, Peru



A dialogue method in a cross-cultural context relies upon the ability of the evaluator to both interact in dialogue with participants, to be able to capture the essence of conversations that take place during group discussion and to assess the quality of the interaction. In this regard the role of interpreters was extremely important. In both Ecuador and Peru the English/Spanish speaking interpreters were highly proficient and appreciated the complexity of the issues being discussed. They helped facilitate the dialogue and put participants at ease. The evaluator also took measures to communicate directly with participants. For example through eye contact or non-verbal responses (eg smiling) while speaking to participants and during their responses. This direct interaction both helped develop rapport and enabled assessment of

the quality of the rapport. In two remote communities (one in Peru and one in Ecuador), secondary translation was required to the local dialect from Spanish. This slowed down the interactions however adequate depth to the exchange was still achieved.

As the first step in the sequence, a diamond was drawn on chart paper and described to participants to be a representation of their community. It was explained that we were going to clarify what 'well-being' meant to them as individuals and to their community. A line was drawn across the centre of the diamond and reported to represent the average case in their community, with the lower and upper extremes representing a high level of well-being and low level of well-being.

The discussion began by identifying the characteristics of the majority of the community or an 'average situation' as the middle of the diamond. Then examples at the extremes of the diagram were identified - for instance what did a high or a low level of 'well-being' mean for the beneficiaries? At times an initial list of prompts was used such as "what makes you and your family happy and content", "what do you aspire to or wish to avoid", "what is important to you in your life", "what do you want in life", "what sort of living conditions are average and better/worse in your community". The list varied depending on the responses. When the group started generating their own ideas, the evaluator would seek similarities and differences within the group and avoid steering the discussion. The dimensions of well-being according to the beneficiaries were verbally expressed and not drawn on the diagram with symbols as has been done in other applications of this method (Mayoux et al. 2005). In general participants were willing to state differences and enter into debate, however participants most often built on and added to thoughts expressed by other participants.

Participants were then asked to individually mark on the chart where they considered themselves to be at two points in time, before and after the World Vision Permaculture project. The chart was either laid on the floor in the centre or passed around the room. The discussion continued while this activity was conducted to shift away attention from where individuals marked their place on the diagram. When all participants had completed their assessment of change due to the project, a discussion explored their reasons for placing themselves in particular areas of the diagram and of attributing changes they had experienced. During this part of the process, the evaluator specifically asked questions to those who had not spoken as much as others. The activity concluded by discussing what it would take to improve their sense of well-being in the future and what makes it possible to move from the lower regions of the diamond to higher up (a greater level of well-being). This revealed greater depth to the meaning of well-being and the sort of future to which they aspired.

The group visual method allowed multiple open-ended questions to be posed around the notion of well-being and a gradual and layered interpretation of the group to form little by little. In addition, the use of both group and individualised attention in the facilitation meant that dialogue involved a range of participants within the group and was rarely dominated by the most articulate or opinionated for more than short lengths of time.

Assessment using the quality framework

Key strengths of the field application

The quality framework highlighted four areas of strength in the field application of this dialogue method:

- The significant contribution to knowledge about beneficiary culture and values that the findings offered and their credibility.
- The depth of the exploration of diversity of perspective.
- The choice and reporting about the intended and actual sample.
- The reflexivity and self-questioning of the evaluator's implicit assumptions.

These strengths provided the potential for the definition of effectiveness to evolve in an appropriate manner, and will now be discussed in turn.

There are two questions in the quality framework which concern how knowledge and understanding has been extended by the evaluation, and the credibility of the findings. Existing knowledge about the lives of the beneficiaries, particularly cultural aspects, was scarce in World Vision Australia, and also limited even in the national office in Ecuador. Instead much of the understanding of the beneficiary communities and their perspective remained with the immediate field staff. The findings were therefore significant in terms of communicating new insights about beneficiaries to other project stakeholders.

The credibility of the findings was ensured through multiple means. Firstly, detailed evidence for the findings was reported in the form of multiple direct quotes from the participants. Secondly, the findings resonated with the understandings of field staff who work closely with the beneficiary communities. The credibility of the findings was also ensured by the fact that participant's spoke in a group environment among their own close-knit community. Finally, a variety of corroborating information collected by staff was used to confirm the findings of the dialogue method, for example through quarterly monitoring reports.

The second strength identified was the depth, richness and diversity of perspective captured, an issue of importance in the quality framework. The repeated questioning of both the group and individuals concerning before and after the Permaculture Project, and concerning aspirations for the future provided a rich tapestry of many different views about people's values, their conceptions of well-being and their hopes. This rich detail and diversity was deliberately maintained in the report by drawing out a large number of themes illustrated with direct quotes from participants. In addition where themes related to a particular community these differences were highlighted, and where the different genders held different views, this was also captured. A further dimension to the diversity of views captured was obtained through a complementary method to the dialogue method. This involved interviews with non-beneficiaries to seek their ideas about well-being. This strategy ensured an alternative perspective was heard, and validated the views expressed by beneficiaries, who had an incentive to portray a positive picture of the project in the evaluation context.

The quality framework requires that the choice of sample and the description of eventual sample coverage be considered. These should be well-justified and well-explicated to enhance transparency of the evaluation process. In collaboration with field staff, purposive sampling sought to capture a diverse set of views. For instance, communities at

different altitudes with different climate and access issues were included. Within particular communities, attendance was hampered by the time of year (harvesting time) and limited availability of beneficiaries. Detailed explanations were provided about these choices and limitations so that the findings could be understood in this context.

Reflexivity is important for a participatory approach and also featured in the quality framework. Reflexivity is a term that describes how the evaluator manages their own place and role in the evaluation process, and the effects they have on data collected and analysis. The evaluator was conscious of being an outsider in a new community, and as someone connected with those who had supplied the funding and implemented the projects. She explained that the purpose of the evaluation was to learn and that sincere responses would be the most valuable for this purpose. She represented herself as an independent party committed to understanding what had been done and how people had been affected. In cases where she was aware that there was a lower level of trust between project staff and community members, she also ensured that project staff were absent from the room during the group discussion. With regard to the method design and analysis, the evaluator reported in detail the steps taken and use of the visual aid in the field to ensure the process followed was transparent.

Limitations of the field application

Applying the quality framework revealed limitations in the use of this dialogue method. These concerned the scope for wider inference and some ethical issues around research procedure.

The quality framework includes a focus on evaluation findings, including the provision of an explanation regarding the scope for wider inference. In other words, a description of what can be generalised (or not) to the wider population from which the sample is drawn and from similar studies. This issue was not discussed in the evaluation report as it was not the intention of the study and therefore not an explicit part of the design. The use of additional complementary methods would have been required to make judgements of this type.

The ethical issues raised by the quality framework are many. They pertain to sensitivity to participants, documentation of the evaluation setting, consent procedures, procedures for protection of confidentiality, protection of anonymity of participants, measures to offer information/advice or services at end of study, and consideration of potential harm or difficulty through participation. Of these, most were considered during the research process. For example a decision was made for consent to be obtained verbally. Consent was acquired by initiating the discussion with an explanation about the evaluation purpose and letting participants know that their participation was totally voluntary and that if there was any question or anything that made them uncomfortable they were urged to say so or to not answer the question. No specific offer was made of information or reporting back at the end of the study. A promise was made by the project staff to the evaluator that the findings would be disseminated and acted upon, however no formal mechanism was set in place to ensure that this happened and no explicit commitment of this was made to the participants. Other proponents of the diagram method report that

diagrams are photographed and then left with participants (Mayoux et al., 2005) and this would have been a positive step. Although the written findings may not have been well disseminated, one positive outcome for the participants was the learning that the participatory method enabled. In some cases this was evidently valued. As one male participant in Lamay, Peru remarked “this discussion lets us think about how we each see ourselves and what we each hope for”. This is in line with the view that a narrator “often discovers meaning while making it” (Gee 1996, p101) and demonstrates the significant value of creating a space for reflection.

Explanation of how effectiveness was defined and refined

Differences between and within each of the beneficiary, implementer and donor perspectives are discussed below, including explanation on how the evaluation process enabled definitions of effectiveness to evolve for some of these groups.

It is important to outline the beneficiary perspective on well-being as a start to this discussion. In Ecuador, one overarching finding regarding conceptions of well-being and important changes which had taken place was evidence of a deep realisation throughout the group of beneficiaries that life can and does change for the better and that this process is in their own hands. Many recognised that the improvements would occur incrementally, as seen during the life of the project. Six common themes arose across the three program areas with regard to changes in their sense of well-being due the project. These related to their level of motivation about life and working on their land, effects on the health of their children, effects on the prospect for the future of their children, happiness that the family unit is able to live together, a change in their sense of connection with their land and their ancestors and a the new found freedom to make choices about their lives. An additional theme prominent in Achupallas the most remote of the communities, was a change in their sense of isolation and hopelessness.

The changes that took place in Peru were in many cases similar to those in Ecuador. For instance similar themes concerning improved life perspective, happiness regarding improvements to the lives of their children and valuing of the recovery of the practices of their ancestors. In addition, beneficiaries also brought up changes related to the improvement of harmony within the family and community structures and the enhancement of the self-esteem and status of women. In one remote community for a few individuals there remained a sense of frustration at problems that seemed unsolvable, for instance access to markets to actually sell their produce. From the beneficiary perspective in both Peru and Ecuador, it becomes apparent that their well-being is very broadly defined and their view of the ‘effectiveness’ of the project was also broad and multi-dimensional.

The project implementer view of ‘well-being’ as a project objective in Peru varied between staff. The field staff as a whole had a significant focus on children, in particular the malnutrition in children. In the same manner as the donor, they defined effectiveness primarily in these terms, though they did also demonstrate an appreciation for a broader definition of well-being. This broader perspective was validated and enhanced through the evaluation process and group discussions in designing the evaluation and in response

to its findings. The Permaculture Project facilitator took a holistic view of the intent of the project from the very start, and gave importance to many different aspects of beneficiary well-being within his approach to the project. His close interaction with the community members (at times facilitated by the remoteness of the communities which meant he stayed overnight in the homes of various families) meant that he had a strong understanding of their values and needs. He also recognised the interconnectedness of different facets of their lives and saw that changes in one aspect affected other aspects. It is likely that the wide-ranging impacts of the project stemmed from his broad interpretation of the project and its purpose. So while his definition of effectiveness may not have changed through the evaluation, it was validated and strengthened.

The project implementer view of the project aims in Ecuador was narrower than in Peru, reflected in the project activities themselves which were more narrowly focused. This narrower view also extended to the expectations about the evaluation and what it would reveal. The project focused on organic farming techniques and irrigation rather than permaculture concepts more widely and this reflected a definition of effectiveness closely linked to food security. In collaborative sessions conducted with staff both before and after the evaluation, it became clear that they were both interested and surprised to learn about the broader impacts of the project that were revealed through the group dialogue method. Their perspective on effectiveness was shifted by this new knowledge and valuing of diverse perspectives.

At the time of commissioning the evaluation, the donor perspective on effectiveness was fairly narrow. From the project documentation, and discussion with key staff, it became clear that the donor conception in World Vision Australia of well-being for this project was primarily focused on reduction of malnutrition within families, particularly children. This view was held within the Program Effectiveness Unit responsible for the commissioning the evaluation, and by the Country Officer. A narrow definition of well-being led to a narrow definition of effectiveness. The open-ended qualitative methods needed to be justified to this audience as the methods broadened the focus beyond what was required for accountability if using these narrow definitions. The final report served to shift and expand the views of the donor and expose the complexity of the social change taking place.

A critical question arises here as to how the process surrounding an actual evaluation method can best be designed to enable and support negotiated interpretations of effectiveness. There are a few approaches which were used to this end in this evaluation that are worth mentioning. The first was using the evaluation process to collect and share the different perspectives of the stakeholders through both the dialogue method just discussed as well as other qualitative methods. The second was the act of building in a commitment to learning into the evaluation. The third was ensuring multiple levels of reporting (that is, both formal reporting through the final report, but also informal discussions and feedback sessions with implementer and donor staff). Finally, a contribution was made by seeking collaboration with the different stakeholders in design and implementation of parts of the evaluation. In retrospect, a further action which might have also been worthwhile was face-to-face debriefing with a group from the donor

office, including both country staff and the Program Effectiveness Unit with an intent to collectively contemplate the implications of the findings.

Implications of the case studies for development practice

The two case studies have demonstrated how definitions of effectiveness can be iteratively refined using narrative and dialogue methods. What becomes apparent through seeing the two case studies side by side, is how an explicit process to facilitate dialogue and debate about effectiveness is embedded in the Most Significant Change monitoring technique. This embedding of the process of negotiation about the meaning of effectiveness is significant, and was not the case for the dialogue method used. For the dialogue method, it becomes clear that such a method needs to be supported by an evaluation terms of reference and design that promotes sharing of perspectives of different stakeholders and engagement in negotiation of the meaning of effectiveness. While a written report has some capacity to catalyse changes in the perspectives of the readers, it may not be read in detail by many of the stakeholders. As such, building in additional processes to support dialogue about effectiveness at different points in evaluation processes with different stakeholders would be a beneficial step towards enabling a continuing evolution and negotiation of meanings of effectiveness.

The critique of the application of both the narrative and dialogue methods using the quality framework demonstrated their strength in capturing and communicating diversity of perspective. This makes them well-suited to the defining and refining of effectiveness. Equally, ethical issues were raised which require on-going attention for the use of qualitative and participatory methods in evaluation. Finally, a need to invest in on-going capacity building in research skills for staff who participate in such methods was demonstrated.

Conclusion

Returning to the ideas expressed at the outset of this chapter, various stakeholders in an aid project are likely to evolve their own definitions of effectiveness over time, and the definition of the term may be negotiated between their different points of view. Narrative and dialogue methods were shown to facilitate this process although both their strengths and their weaknesses must be acknowledged. The value in monitoring and evaluation approaches that support negotiation around effectiveness is that they enable accountability to beneficiaries, implementers and donor staff. In addition, through participation in such methods, or exposure to the findings they produce, stakeholders are able to reach an improved level of personal reflexivity about their role and their values. In particular this process allows donors and implementers to better align their actions with the perspectives of the beneficiaries they seek to serve. The ultimate outcome is more valuable, appropriate, and valued development aid.

References

Barry, M. & Fourie, C. 2001, 'Wicked Problems, Soft Systems and Cadastral Systems in Periods of Uncertainty', *Proceedings of FIG International Conference on Spatial*

- Information for Sustainable Development*, Department of Geomatics, University of Cape Town South Africa, Nairobi, Kenya.
- Chambers, R. 1997, *Whose reality counts? Putting the last first*, Intermediate Technology Publications, London.
- Checkland, P. 2001, 'Soft Systems Methodology', in J. Rosenhead & J. Mingers (eds), *Rational Analysis for a Problematic World Revisited*, 2nd edn, John Wiley & Sons Ltd, Brisbane.
- Crawford, P. 2004, 'Aiding aid: a monitoring and evaluation framework to enhance international aid effectiveness', University of Technology, Sydney.
- Fowler, A. 1996, 'Assessing NGO Performance: Difficulties, Dilemmas and a Way Ahead', in M. Edwards & D. Hulme (eds), *Beyond the Magic Bullet: NGO performance and accountability in the post-cold war world*, Kumarian Press Inc., West Harford.
- Gee, J. P. (1996). *Social Linguistics and literacies: Ideology in Discourses. 2nd Edition*. London: Taylor & Francis Ltd
- Kaplan, A. 1999, *The developing of capacity*, Community Development Resource Association.
- Kelly, L. & Chapman, R. 2003, 'A process to define and enhance non-government organisation effectiveness', *Australian Project Manager*, vol. 23, no. 2, pp. 29-31.
- Madeley, J. 1991, *When Aid is No Help: How projects fail, and how they could succeed*, Intermediate Technology Publications, London.
- Mayoux L. and Chambers R. (2005) 'Reversing the Paradigm: Quantification, Participatory Methods and Pro-Poor Impact Assessment', *Journal of International Development* Vol 17, p271-298
- Mayoux L, Andharia J, Dand S., Hardikar, N. and Thacker, S. (2005) 'Participatory action learning system (PALS): Impact assessment for civil society development and grassroots-based advocacy in Anandi, India', *Journal of International Development* Vol 17, p211-242
- Reissman, CK (1993) *Narrative Analysis (Qualitative Research Methods Series 30)* Newbury Park, CA : Sage Publications
- Smutylo, T. 2001, 'Crouching impact, hidden attribution: overcoming threats to learning in development programs', *Workshop on Across Portfolio Learning*, Evaluation Unit, International Development Research Centre, Ottawa Canada, Block Island, <http://www.idrc.ca/evaluation/crouching_impact.pdf>.
- Wadsworth (1984) *Do it yourself social research*, Allen and Unwin, Crows Nest Australia
- Willetts, J. 2004, *Evaluation of the Pilot Project Most Significant Change by ADRA Laos*, Institute for Sustainable Futures, University of Technology, Sydney Available at <http://www.isf.uts.edu.au/publications/willetts.html> (retrieved May 2007).
- Willetts, J. and Crawford, P. 2007, 'The Most Significant Lessons about the Most Significant Change Technique', *Journal of Development Practice*, 17(3), p367-379

Juliet Willetts is a Researcher Principal at the Institute for Sustainable Futures, University of Technology (www.isf.uts.edu.au). She specialises in transdisciplinary research that supports improvements to policy and practice in international development aid and sustainable water management. Her background includes a doctorate in Environmental Engineering from University of NSW, an extended period of in-country work in health education in India, and numerous research consultancies in Australia and internationally.

Helen Cheney has worked as a researcher, educator and manager in community, academic and government organisations, including roles at the Centre for Appropriate Technology, the Centre for Aboriginal Studies, Curtin University of Technology and the Institute for Sustainable Futures, University of Technology, Sydney. Helen is currently a Senior Research Fellow at the Australian Institute of Family Studies. Her fields of expertise are in community development and participatory evaluation.

Paul Crawford is a Monitoring and Evaluation (M&E) specialist with Aid-IT Solutions (www.aid-it.com.au) and an Adjunct Research Fellow at the Institute for Sustainable Futures where he completed his doctorate in M&E. Paul has consulted to a variety of bilateral, multilateral, NGO and contractor agencies throughout Asia, Africa, the Pacific, the Middle East, the Balkans and the former Soviet Union. He continues to be intrigued by how aid effectiveness can be enhanced.