

A Dialogical, Story-Based Evaluation Tool: The Most Significant Change Technique

JESSICA DART AND RICK DAVIES

Dr Jessica Jane Dart

Evaluation Support Team, Division of Agriculture, Department of Primary Industries,
Hazeldean Road, Ellinbank, RMB 2460, Warragul, Victoria 3820, Australia

Jessica.Dart@nre.vic.gov.au or Jessdart@austarmetro.com.au

Tel (+61 3) 5624 2222, (+ 61 0425 735530)

Fax (+61 3) 5624 2200

Dr Richard Jonathan Davies

82 Bishops Rd, Trumpington, Cambridge CB2 2NH, United Kingdom

rick@shimbir.demon.co.uk

Pre-print copy: Copyright of American Journal of Evaluation

For use of workshop participants only

Reference

Dart, J. J. & Davies R.J. (forthcoming) A dialogical story-based evaluation tool: the most significant change technique, *American Journal of Evaluation*.

ABSTRACT

The Most Significant Change (MSC) technique is a dialogical, story-based technique. Its primary purpose is to facilitate program improvement by focusing the direction of work towards explicitly valued directions and away from less valued directions. MSC can also make a contribution to summative evaluation through both its process and its outputs. The technique involves a form of continuous values inquiry whereby designated groups of stakeholders search for significant program outcomes and then deliberate on the value of these outcomes in a systematic and transparent manner. To date, MSC has largely been used for the evaluation of international development programs, after having been initially developed for the evaluation of a social development program in Bangladesh (Davies, 1996). This article introduces MSC to the readers of this journal and discusses its potential to add to the basket of choices for evaluating programs in developed economies. We provide an Australian case study and outline some of the strengths and weaknesses of the technique. We conclude that MSC can make an important contribution to evaluation practice. Its unusual methodology and outcomes make it ideal for use in combination with other techniques and approaches.

INTRODUCTION

The Most Significant Change technique (MSC) was invented to meet some of the challenges associated with evaluating a complex, participatory, rural development program in Bangladesh (Davies, 1996) and is now used by many international development organizations. It represents a radical departure from the conventional monitoring against quantitative indicators that is commonly seen in this sector. MSC involves the regular collection and participatory interpretation of “stories” about change rather than predetermined quantitative indicators. The aim of this paper is to introduce MSC to practitioners in the field of program evaluation. To this end, we provide a basic overview of MSC and present a case study of its implementation in a developed economy. We attempt to locate it within the wider discourse of program evaluation and explore its potential to add to the range of options available for evaluating programs. MSC has also been referred to as 'the Evolutionary Approach to Organisational Learning' (Davies, 1996), the 'Story Approach' (Dart, 1999a, 1999b,) and 'monitoring without indicators' (Guijt et al., 1998). We now prefer the term 'Most Significant Change', because we believe that the central aspect of the technique is not the stories themselves, but the deliberation and dialogue that surrounds the process of selecting significant changes.

Overview of MSC Methodology

MSC has seven key steps (Davies, 1996):

1. The selection of domains of change to be monitored
2. The reporting period
3. The participants
4. Phrasing the question
5. The structure of participation
6. Feedback
7. Verification.

Firstly, the people managing the MSC process identify the domains of change they think need to be evaluated. This involves selected stakeholders identifying broad domains—for example, ‘changes in people’s lives’—that are not precisely defined like performance indicators, but are deliberately left loose, to be defined by the actual users.

Stories of significant change are collected from those most directly involved, such as beneficiaries, clients and field staff. The stories are collected with the help of a simple question: ‘During the last month, in your opinion, what was the most significant change that took place in the program?’ It is initially up to respondents to allocate their stories to a domain category. In addition to this, respondents are encouraged to report why they consider a particular change to be the most significant one.

The stories are then analyzed and filtered up through the levels of authority typically found within an organization or program. Figure 1 illustrates the flow of stories and feedback. Each level of the hierarchy (in this case: regional committees, state committee and funders) reviews a series of stories sent to them by the level below and selects the single most significant account of change within each of the domains. Each committee sends the ‘winning’ stories up to the next level of the program hierarchy, and the number of stories is whittled down through a systematic and transparent

process. Every time stories are selected, the criteria used to select them are recorded and fed back to all interested stakeholders, so that each subsequent round of the story collection and selection process is informed by feedback from previous rounds. The organization is effectively recording and adjusting the direction of its attention—and the criteria it uses for valuing the events it sees there.

At the end of each period, such as a year, a document is produced with all the stories selected at the uppermost organizational level over that period. The stories are accompanied by the reasons the stories were selected. This document contains several chapters with the stories selected from each of the domains of change. It is forwarded to the program funders and they are asked to assess the stories, selecting those that most fully represent the sort of outcomes they wish to fund. They are also asked to document the reasons for their choice. This information is fed back to program managers.

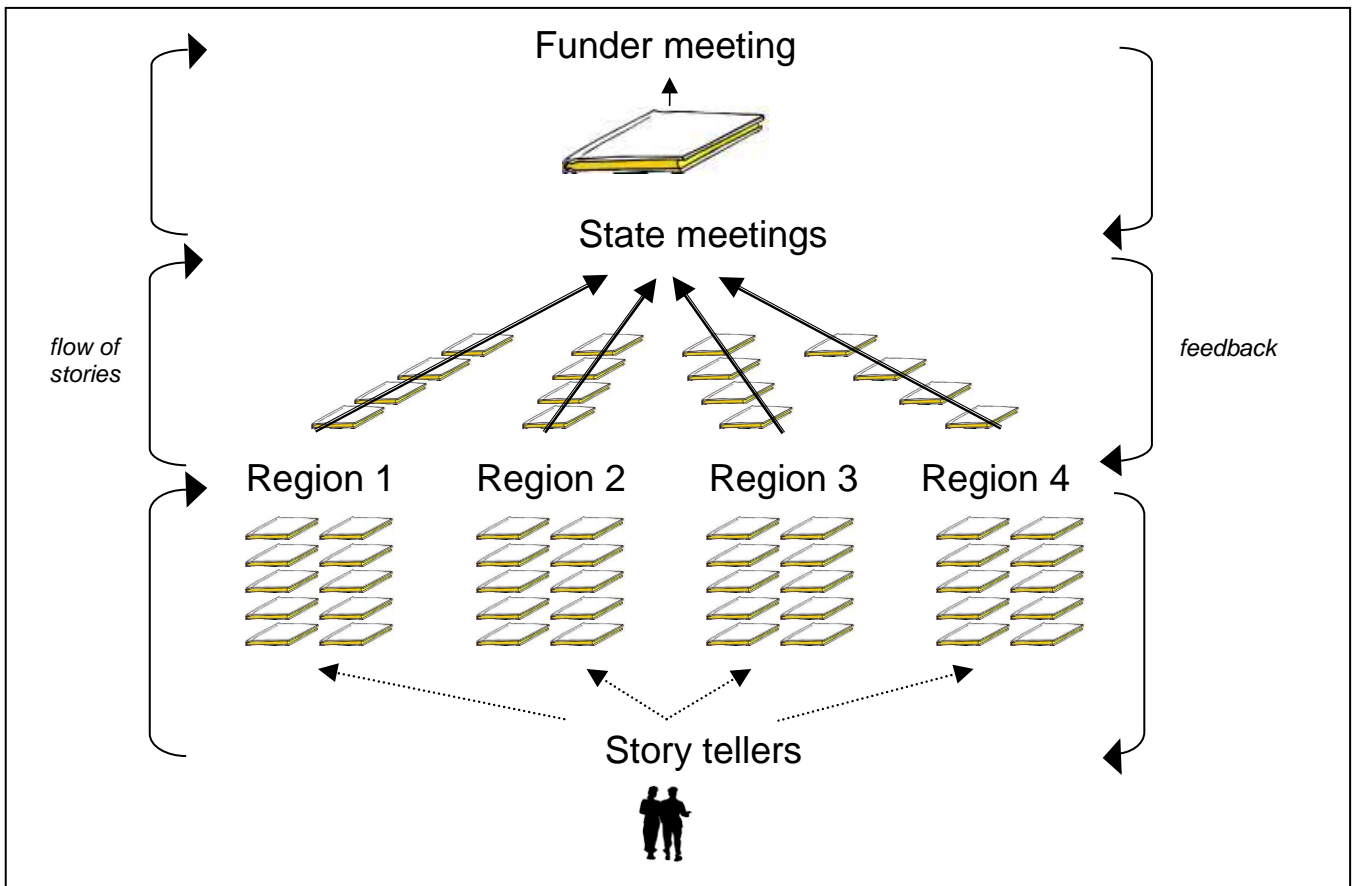


Figure 1. Flow of Stories and Feedback in MSC.

The winning stories can then be verified by visiting the sites of the described events. The purpose of this is twofold: to check that storytellers are reporting accurately and honestly and to provide an opportunity to gather more detailed information about events seen as specially significant. If conducted some time after the event, the visit also offers a chance to see what has happened since the event was first documented.

In addition to the seven steps outlined above, there are two optional steps (Davies, 1996). The optional eighth step is quantification, which can take place at two stages. When an account of change is first described, it is possible to include quantitative information as well as qualitative information. It is also possible to quantify the extent to which the most significant changes identified in one location or zone have taken place in other locations, within a specific period. The

optional ninth step is monitoring the monitoring system itself, which can include looking at who participated and how they affected the contents, and analyzing how often different types of changes are reported.

MSC: A POWERFUL EVALUATION TOOL

Purpose of MSC

Patton (1997, p65) suggests that evaluation findings can serve three primary purposes: ‘rendering judgments, facilitating improvements and/or generating knowledge’. Against this framework, the key purpose of MSC is facilitating program improvement. This is facilitated by focusing the direction of work towards explicitly valued directions and away from less valued directions.

MSC also contributes important information and processes to help judge programs (summative evaluation), but it is not intended as a stand-alone technique. Instead, MSC acts as a ‘satellite’ tool, providing a series of functions to enhance the core summative evaluation effort. These functions include: the provision of information concerning unexpected outcomes; performance information by way of the best success stories and a form of dynamic values inquiry. We discuss these functions more fully after the case study.

Underlying mechanisms: How MSC works

Dynamic values inquiry. MSC can be conceived as a form of dynamic values inquiry whereby designated groups of stakeholders continuously search for significant program outcomes and then deliberate on the value of these outcomes. This process contributes to both program improvement and judgment. Of late there has been discussion of the importance of values and values inquiry in evaluation (Henry, 2002; Mark, et al, 2000; House & Howe 1999). Advocates suggest that criteria used to judge programs do not always reflect stakeholder values and that uncovering these values can help to ensure that programs meet needs.

Dynamic values inquiry is a central and critical part of MSC. When key stakeholders select stories of significant change, they partake in an ongoing process of deliberation about the value of individual outcomes. As MSC occurs over time, the dialogue is responsive to the changing nature of the program and its context. The process of values inquiry occurs throughout the organizational hierarchy, ultimately involving program funders and policy makers. These deliberations about the value of outcomes contribute to program improvement by facilitating an ongoing, organization-wide conversation that guides the direction of work towards explicitly valued directions.

In MSC, program staff, stakeholder committee members and funders all spend considerable time deliberating how individual client/beneficiary outcomes should be judged, and their chosen criteria are documented along with the stories. This repeated deliberation concerning the different criteria used to judge outcomes can contribute to summative evaluation efforts in both tangible and intangible ways. Evaluators can examine the stories and interpretations to get a good understanding of the criteria used by MSC participants. On interviewing MSC participants, evaluators would also expect to find that participants have a clear picture of what outcomes they value, and what outcomes other stakeholders value. This information concerning stakeholder values could be used in planning a summative evaluation in a similar way as the sort of values inquiry described by Mark et al., (2000) however an investigation of additional stakeholder views may also be necessary.

Stories. Stories are a valuable part of MSC for several reasons: they encourage non-evaluation experts to participate, they are likely to be remembered as a complex whole, and they can help keep dialogue based on concrete outcomes rather than abstract indicators. Storytelling is an ancient and cross-cultural sense-making process familiar to all peoples. The stories make MSC more human, and people seem to relate to the information more when it is told in the story format.

Shaw et al. (1998) contend that stories are central to human intelligence and memory. A good story defines relationships, a sequence of events, cause and effect, and a priority among items—and those elements are likely to be remembered as a complex whole. Stories about the

impact of interventions can infiltrate the collective memory of an organization, helping program staff to gain and retain a more deeply shared understanding of what is being achieved. This creates a common base to enter into dialogue about what is desirable in terms of expected and unexpected outcomes. Boje (1991) contends that, in complex organizations, part of the reason for storytelling (in casual conversation) is the working out of value differences at the interface of individual and collective memory. In an evaluation context, stories are an ideal medium for practitioners to make sense of the array of program outcomes and stakeholder values.

Despite disagreement across the literature about the precise definition, stories are frequently defined by their structure, being treated as discrete units with clear beginnings and endings. In an evaluation context, stories emerge during interviews (often embedded in transcripts) and in written documents such as diaries or open-ended responses to questions. They differ from other forms of interview response in that they are detachable from the surrounding discourse rather than being situated events (Riessman, 1993). Thus they often tend to emerge when the teller is left to complete their narrative without interruption. In MSC, the term 'story' refers to an account of change told in response to a specific question (i.e., 'what was the most significant change that occurred for you in the last month as a result of the program?'). However, the teller is also asked to provide some descriptive and interpretive information after the story has been narrated. In the case study given in this article, the stories tended to be one to two pages in length (for examples see Boxes 2 and 3).

How MSC Relates to Other Techniques and Approaches

MSC has points in common with the critical incident technique (CIT) often cited in literature from the health sector (e.g., Jacobs et al., 1978; Starr-Schneidkraut et al., 1997). Both MSC and CIT ask stakeholders to recall memorable experiences of what they consider to be critical (significant) events. Both techniques can also involve ongoing reporting of events or incidents deemed to be significant by stakeholders. A key difference between them is that CIT focuses on variations from prescribed practice and generates negative information, whereas MSC searches for significant outcomes through an inductive process, and tends to generate mainly positive information.

MSC also has much in common with Kibel's results mapping (1999), which makes use of client stories to assess performance. Kibel's approach, developed to help evaluate programs engaged in healing, transformation and prevention, differs from MSC in that stories are coded by experts against a results ladder and a contribution analysis. MSC stories are filtered up through the organization, through a participatory process involving values inquiry that is systematic and transparent. Nonetheless, both processes attempt to solve a fundamental issue: how to do justice to the diversity and complexity of individual client/beneficiary outcomes.

MSC Summarized

MSC is a dialogical, story-based technique. Its primary purpose in evaluation is to improve the program by focusing the direction of work towards explicitly valued directions. However, it can also contribute to summative evaluation through values inquiry and the provision of information concerning unexpected and most successful client outcomes. The underlying mechanism is a form of continuous values inquiry whereby designated groups of stakeholders search for significant program outcomes and then deliberate on the value of these outcomes. It has points in common with critical incident technique and Kibel's results mapping (1999).

CASE STUDY: IMPLEMENTING MSC IN AUSTRALIA

Evaluation research was conducted between 1998 and 2000 (Dart, 2000) to investigate the potential for using MSC in evaluating public sector programs in developed economies. The work involved implementing and adapting MSC for an agricultural extension program in Australia. The purpose of using MSC in the case study evaluation was to supplement the existing evaluation portfolio and to promote organizational learning.

Background to the Case Study Program

Target 10 is a collaborative dairy extension program that works with farmers to improve farm productivity and profitability in a sustainable manner. The program focuses on issues of high

priority to the industry such as grazing management, business management, dairy cow nutrition, soils and fertilizers and natural resource management. Information is extended to farmers through courses, discussion groups, newsletters, comparative analysis, field days, focus farms, demonstrations and other activities. In 1998, the program employed around 50 staff and operated across four regions of the Australian State of Victoria.

Since its inception in 1992, Target 10 had used a wide variety of evaluation techniques, including economic benefit-cost analyses, comparison group trials, and market segmentation (McDonald et al. 2003). At the time of this case study, the program team were also beginning to develop a program theory approach using Bennett's Hierarchy – a hierarchy of goals established for agricultural extension programs (Bennett, 1975; Bennett & Rockwell, 2000). However, in 1997, stakeholders still felt that certain areas of evaluation needed to be developed. In particular, they wished to understand more about the impact of the program on farmers' lives.

The organizational structure under which the program operated was complex, having both public and private stakeholders and partnerships with a university and the dairy industry. The program also had stakeholder steering committees at regional and state levels. These committees were chaired by farmers and comprised farmer representatives, extension staff, university staff and representatives from the local dairy industry. Taking this organizational complexity into account, it was vitally important that time be allocated for the various stakeholders to enter into a meaningful dialogue about what was happening in the field, and whether these experiences represented the sort of outcomes that were desirable.

MSC was chosen because it appeared to fit well with the various evaluation needs given the existing evaluation portfolio and program context, and it was agreed that the approach be implemented across the whole program for one year.

Implementation of MSC across the Case Program

In the Target 10 case, MSC was implemented in six research phases:

1. Pilot test and become familiar with the process
2. Establish the domains of change
3. Establish a reference group
4. Establish a process to collect and review stories of change
5. Hold an annual round-table meeting for program funders to review the stories
6. Conduct a secondary analysis of all stories.

1. **Pilot test and become familiar with the process.** During the early stages of introducing the approach (May 1998), it became apparent that staff and committee members were not sure what sorts of stories were required. They asked for guidance on the length of stories, suitable subject matter and the form stories should take. Staff were understandably nervous about putting pen to paper, so a form was developed (see Box 1) to help collect the stories. The approach was pilot tested by the program team in a pre-trial workshop in June 1998.

Story title:	'.....'
Domain:	<input type="checkbox"/> changes in decision-making skills <input type="checkbox"/> changes in on-farm practice <input type="checkbox"/> changes in profitability/ productivity <input type="checkbox"/> other significant changes
Name of person recording story:
Region:
Date of narration:	... /... /.....
Where did this happen?
When did it happen?

What happened?	
Why do you think this is a significant change?	
What difference has it made already/ will it make in the future?	

Box 1. Types of Information Requested in the Story-Collection Form (*the actual form allocates more space for responses*)

2. **Establish the domains of change.** Domains of change are loose categories of change used to distinguish different types of stories. In the case of Target 10, the domains of change were

established using the Delphi techniqueⁱ (Delbecq et al., 1975) and involving more than 150 program stakeholders. The domains of change chosen for monitoring the Target 10 dairy extension program were: changes in on-farm practice, changes in profitability or productivity, changes in farmer decision-making skills and any other significant types of change.

3. **Establish a reference group.** A reference group was established in September 1998 to capture learning, to encourage the adaptation of the process to local conditions and to coordinate the process. The reference group comprised four extension staff, representing each of the four program regions, and the facilitator. These staff had volunteered to help coordinate story collection and selection in their regions, and were referred to as the ‘story champions’. Any modifications to the process were discussed with these four staff. In some cases, an idea was tested in one region before recommending it to other regions.

4. **Establish a process to collect and review stories of change.** In June 1998, all staff and committee members (farmers, extension staff, university staff and local industry representatives) were supplied with blank proformas and asked to generate stories concerning what they considered to be significant changes (see Box 1). However, as this method produced very few stories, stakeholders were encouraged to share stories verbally during the meetings. These impromptu stories were recorded and later transcribed. For some people, this was their preferred form of storytelling.

It was decided at an early stage of the implementation that MSC should ride on the back of the existing program structure. This was because stakeholders did not want any additional meetings. The stories were to be collected primarily by staff and stakeholder members of the steering committees, based on their own experience or second-hand accounts from farmers and other stakeholders. The storytellers would nominate the appropriate domain for each story and indicate why they thought that story was significant. Each of the four regional committees

would then select one story from each of the four domains to send to the state-wide committee, which would select a further four stories at each meeting to be sent to an annual round-table meeting with the program funders.

Each region developed its own system of selecting and collecting stories, depending on the existing committee structure. The only condition was that all regional committees should submit four stories to each state meeting (held every two or three months) and that they should document how they had selected these stories.

The stories were selected at meetings with the help of a facilitator. Story titles were recorded under the respective domains. When all the stories had been read out, each domain was considered in turn. The facilitator then asked questions to encourage debate before moving on to a vote by show of hands. Each committee member was given one vote for each domain. If there was no consensus, which was common, further discussion was facilitated until the meeting could agree on which story should be selected. Agreement was normally achieved by an iterative voting and discussion process. If no agreement could be reached, either two stories were selected, or none. Some stories were selected with the proviso that a caveat be attached to the story explaining additional factors, or indicating that not all committee members agreed about the value of the outcome. The idea was to come to an agreement as a group. As well as selecting a story, committee members were also asked to state why that story had been selected above the others. Much of the discussion revolved around explanations of why committee members thought one story was particularly valuable or misleading. This discussion was recorded on tape or by a note-taker.

It was during this phase that the values of various stakeholders were revealed. The 'game' of voting for one story above the others invoked deep discussion about the value of the outcomes. Stakeholders consistently began with differing opinions of the value of the outcomes represented in the stories. The opinions of the farmer representatives were given particular

regard during these discussions, because the steering committees had been established to ensure that the program was farmer-driven.

Boxes 2 and 3 provide examples of stories selected by steering committees. Each story is accompanied by a brief explanation of why it was selected. Box 2 gives an example of a success story with expected client outcomes, while Box 3 presents unexpected client outcomes.

Title	Knowledge is Power
Name of person recording story	Dairy farm employee
Region	XXX
Date of narration	Round 5 – 15 March 1999
Where did this happen	XXX
When did it happen	1996-1999
<p>What happened? <i>In 1996, I was working on a 500-cow dairy farm and had no input to the day-to-day running. But I wanted to have more input. I decided to do something about it, so I attended the Target 10 Grazing Management Program followed by the Nutrition Program. After completing these courses, I had enough knowledge to start making decisions, such as measuring growth rates and adjusting rotation lengths. Gaining this knowledge led me to start a new job on a 550-cow dairy farm where I am presently managing all feed requirements.</i></p>	
<p><i>I have been able to maximise pasture consumption to 10 tonnes per hectare; and, with 1 to 2 tonnes of grain fed, this will achieve a production of 1,000 kg of solids per hectare on 850 mm of rainfall per year. I walk the farm weekly to measure growth rates, allowing me to adjust the rotation to suit the growth rate. I fill the gap with grain so as not to waste pasture. When harvest comes, I can use this method to cut maximum fodder without sacrificing the cows' pasture.</i></p>	
<p>Why do you think this is a significant change? <i>These changes, which I have been able to apply to the day-to-day running of the farm, have made the farm more efficient. I am also a more efficient employee now, and I have seen what I can achieve by furthering my studies. Last year I completed a Diploma of Agriculture (dairy farm management), and this year I'm studying an Advanced Diploma (dairy farm management). My future plan is to manage the whole day-to-day running of a larger scale dairy farm in every aspect.</i></p>	
<p>Feedback from the central executive committee:</p> <ul style="list-style-type: none"> • <i>This story is a good example of one person going on to do more learning and expand his horizons.</i> • <i>It is good, as it is written by a farmer.</i> • <i>It is about building confidence, and the story even got down to changes in productivity.</i> • <i>It really shows the full picture.</i> 	
<p>Feedback from the round-table meeting:</p> <ul style="list-style-type: none"> • <i>Good positive story about practical learning.</i> • <i>Real change directly related to having done Target 10 programs (they look good in the résumé) with measurable results as well.</i> • <i>Substantial behaviour change as a result of the Target 10 program. But the profit/productivity impact is not so clear.</i> • <i>Great story. Self-improvement outcomes, the subject has strong commitment to the industry and the desire to achieve. These are the people that will make the industry move forward.</i> • <i>To me it is raw efficiency; that's what really gets me. It is almost the most you would expect to get from a program. The guy is really ready to roll and he has got the right attitude.</i> 	

Box 2. Example of a Success Story.

Title	Family Starts to Communicate
Name of person recording story	XX Target 10 extension officer
Region	South West
Date of narration	Round 3 – 7 October 1998
Who was involved	XXX
When did it happen	XX

What happened? During a Dairy Business Focus workshop, a female participant become upset about the way her family business operated. She, her husband and her parents-in-law all worked on the same property. She was upset because she had no idea of their financial position, and her immediate family had a very small living allowance. She claimed that, after 20 years of living like this, of having no financial control, let alone any idea of the financial position of the farm and being left out of strategic decision-making, she had had enough. (She told a story of how she thought the farm had no money, when her parents-in-law bought the farm next door and virtually paid cash for it, with no consultation with her or her husband.) I tried to encourage her to contact a rural counsellor, as did other participants, because she was very upset. I then spoke to a rural counsellor and told her about this farmer's situation and invited the rural counsellor to the next meeting. To my surprise, the farmer did not speak to the rural counsellor at that meeting.

To my further surprise, this troubled dairy farmer submitted her Dairy Farm Performance Analysis (DFPA) input form with some other participants. At this next workshop, I then approached this farmer and asked why she hadn't spoken to the rural counsellor. She said that she had initially been frightened to stir up trouble and make the situation worse. She then explained that, to complete the DFPA input document, she had to call a family meeting and get her husband and parents-in-law to help complete the sheet, as she had no idea about the financial situation of the farm. At that meeting, she was amazed at how willing her parents-in-law were to explain the costs and income. She began to see that her parents-in-law were prepared to share information with her and her husband, but that she had never asked them to!

We then talked about trying to arrange regular family meetings and that, if her parents-in-law thought they were a waste of time, to find excuses, such as filling in the DFPA form, to hold such meetings. She realised that she didn't need to threaten to leave if things didn't change and that, by being pro-active, she could solve her own problem.

Why do you think this is a significant change? For the past 20 years, this farmer had been unhappy with her situation. After a three-day workshop and given the practical task of having to complete a physical and financial DFPA input form, the family has started to communicate.

What difference has it made already or will it make in the future? The farming relationship with her parents-in-law had got to point where she could tolerate it no longer. This relationship has now become more sustainable/tolerable and, with improved communication, not only should all concerned be happier, but this may also lead to changed farm management practices and an increase in profit.

Feedback from the Central Executive Committee:

- There is lots of emotion in this story. It has a very personal approach. It was a gripping story. It had a beginning, middle, and an end! People reading this story got involved in the story, and it was well written.**
- The lady in this story has found out that she can communicate! This is about a change in communication. But it is also about potential change in a lot of other things.**

Box 3. Example of an Unexpected Outcome.

5. Hold an annual round-table meeting. At the end of the 12-month trial period, a round-table meeting was held with eight participants who were considered to be funders or ‘key influencers’ of the Target 10 program. The meeting took the form of a facilitated group discussion in which participants were asked for their reaction to 24 stories. Prior to the meeting, the funders had been asked to score the stories individually. The scores were presented to the group, followed by a discussion as to which outcomes in the stories were valued and for what reasons.

When the scores were examined, it became apparent that the funders had very different reactions to the stories. Two stories were allocated the lowest score by one funder and the highest score by another! The group did not have a unified vision as to what was desirable. It was also apparent that most of the stories were considered to have merit by at least one of the participants. After some discussion, the group eventually agreed on one story (see Box 2) that represented the sort of outcomes they all wanted to fund. (It is interesting to note that the program is now delivering more of this type of outcome.) The group’s discussion of all 24 stories was recorded and a summary of the funders’ comments provided to program staff. The staff felt somewhat vindicated when they discovered there was little agreement between the various funders because they had already suspected that different funders valued different program outcomes.

In addition to circulating feedback from the statewide committee, a booklet with all the stories selected by the statewide committee during the year was distributed to program stakeholders. Each story was accompanied by the storyteller’s interpretation and comments from the statewide committee and the funders. Thus a reader could judge the story themselves, and also find out how the story was valued by the program’s committees and funders.

6. Conduct a secondary analysis of all stories. In total, 134 stories were collected, transcribed and entered into a database over a 12-month period. Of these stories, 77 percent were generated by program extension staff, 13 percent directly by farmers and 10 percent by industry and

university representatives. The stories were also analyzed en masse, and the findings included in the booklet circulated to all stakeholders. This analysis involved examining the origin of the stories, the main themes and the differences between the stories that were selected and those that were not. The outcomes represented in the stories were also matched to different levels of a program logic model structured against Bennett's Hierarchy (Bennett, 1975), and it became clear that participants had shown a preference for stories relating to higher levels of the outcome hierarchy. Our experience in development and extension programs is that outcomes at these higher levels are often the hardest to document.

Troubleshooting the process.

It would be misleading to suggest that MSC was implemented smoothly and easily across the program. At various stages in the 12-month trial, problems arose and were addressed where possible. As the process was an iterative one, it was possible to modify each round on the basis of feedback from the previous round. The main problems were associated with the time taken to run the process and the need to develop a system to ensure confidentiality. Some people also disliked the competitive aspect of the process, feeling disillusioned when their stories didn't get selected. It was also noticeable that the program's four regions responded differently to MSC.

About 10 percent of all stories collected involved an element of 'bad news'. Feedback from the statewide committee suggested that it was beneficial to read and discuss stories associated with negative outcomes. Because of this, eight months into the process, a fifth domain of change was added: lessons learned. Including this as a domain implied that each region was obliged to submit one story about lessons learned for every statewide review.

THE VALUE OF MSC FOR THE CASE STUDY

After the 12-month trial, we examined the extent to which MSC contributed to the purposes for which it was intended and the extent to which it represented a valid tool for inquiry, exploring the following questions:

1. To what extent did MSC facilitate program improvement?
2. To what extent did MSC contribute to summative evaluation?
3. To what extent did the processes encourage stakeholder participation?
4. To what extent were the design, data collection and analysis valid?

Our research was based on empirical evidence that included transcripts from all review sessions, semi-structured interviews with key stakeholders, a group interview with the funder panel, and an internet questionnaire completed by all program staff. Key informants also participated in the development of a model showing how MSC was expected to bring about program improvements (see Figure 3). While it is beyond the scope of this paper to present this particular work in detail, the main results are outlined below. A more detailed description of the research is available elsewhere (Dart, 2000), as is an analysis of the value of the process against an organizational learning theory framework (Davies, 1998).

1. To What Extent Did MSC Facilitate Program Improvement?

In MSC, program improvement is facilitated by focusing the direction of work towards explicitly valued directions and away from less valued directions. MSC was originally framed as a tool to enable organizational learning (Davies, 1996). This can be done by making an organization's existing processes of observation and judgment more visible, and thus more open to change. This representation of the process of organizational learning is based on evolutionary epistemology,

which uses the Darwinian theory of evolution, adaptation, and natural selection to describe organizational change (Campbell, 1969; Nelson & Winter, 1982). Within the evolutionary view, prior organizational learning is embodied in organizational routines, which are sometimes modified in the face of changing internal and external environments. Some changes have more ‘fit’ than others, favoring the survival of those routines, their users and the organization. The idea as applied to MSC is that the fittest (deemed most significant) reported changes are identified through a selection process embodied in the design of the evaluation system. Because these are endorsed as the most desirable outcomes, the whole organization can then create more of these explicitly valued outcomes.

Because changes in organizational learning were expected to be difficult to assess after just 12 months, which was the time allocated to the case study, a model was constructed early in the trial year of how the program staff felt the approach could bring about program improvement. The model was developed through an iterative process of consultation with program staff. Figure 2 summarizes the model, which was organized with the help of Bennett’s Hierarchy (Bennett, 1975). Patton, (1997, p236) uses the same hierarchy to structure an evaluation theory model for utilization focused evaluation.

The intermediate outcomes presented in the model (changes in behavior, knowledge, reactions and participation) were tested using empirical data after the one-year trial was complete. The extent to which these intermediate outcomes were perceived to have been achieved varied considerably with program contexts. Nonetheless, the model depicted in Figure 2 remains a fair representation of how respondents felt MSC led to program improvements.

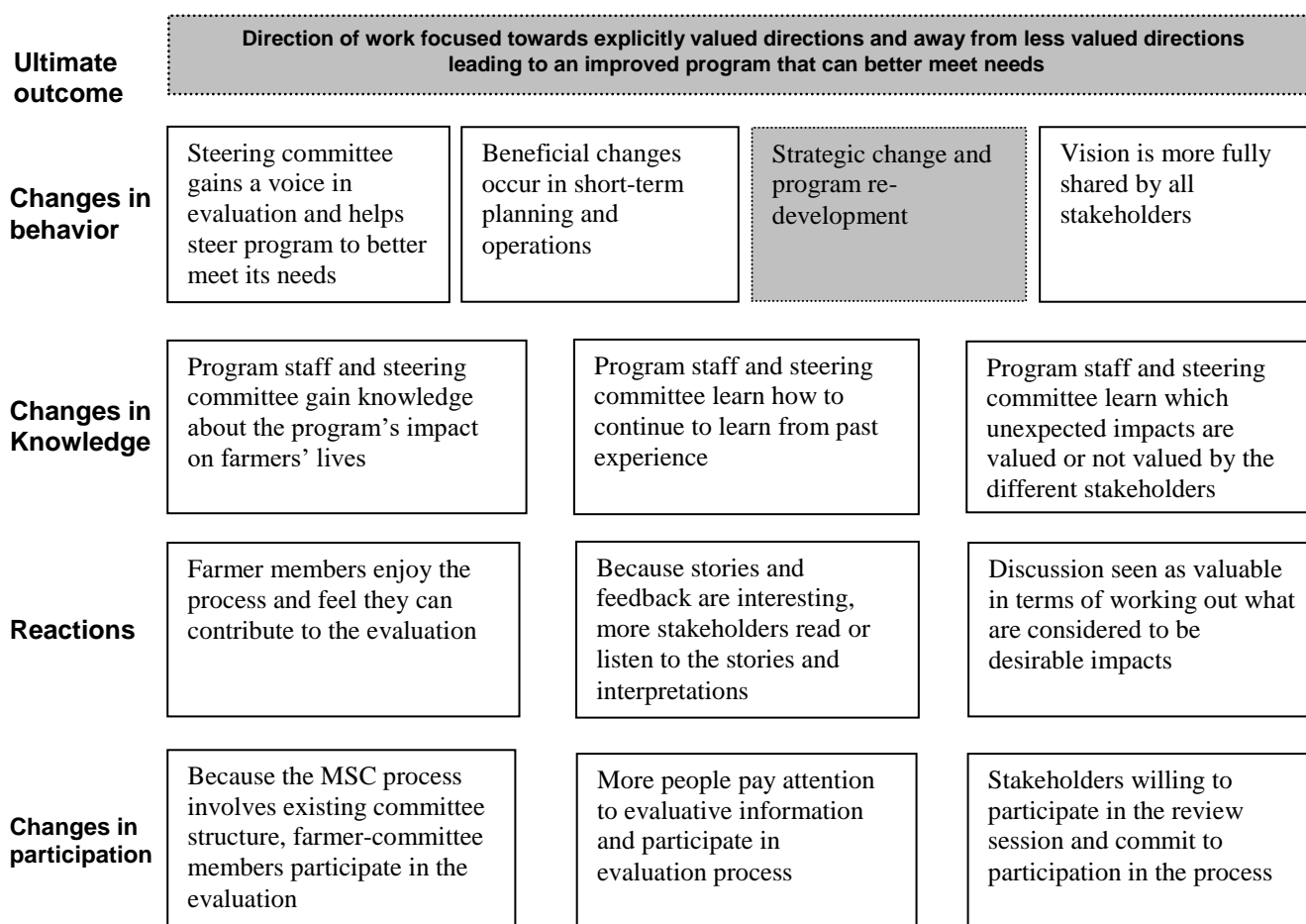


Figure 2. An evaluation theory model showing how MSC was thought to bring about program improvements. (The shaded box indicates that informants felt it was not yet clear whether MSC had influenced these outcomes.)

A survey of staff and key stakeholders indicated that the farmer committee members particularly enjoyed MSC, and some felt this helped farmers to have a voice in the evaluation. Respondents also suggested that MSC helped them to understand the program impact's on farmers' lives and to appreciate what outcomes were valued by different stakeholders. Respondents felt that, in the long term, the MSC process could help program stakeholders reach a more fully shared vision. Many program staff reported that they had used the stories and interpretations to help them plan new activities, with stories of lessons learned seen as particularly useful in this regard.

However, some respondents felt that MSC lacked a system for responding to the lessons gained from the stories.

There was less evidence for the stories' influence on the program redevelopment in terms of large-scale strategic changes, some of which had occurred shortly after the 12-month trial (presented as shaded box in Figure 2). Informants inferred that the stories did not directly affect operational and strategic changes because these were influenced by many other factors.

Respondents also identified some benefits not predicted at the beginning of the trial or included in the model. Firstly, more than half the program staff explained how they had used MSC stories in their normal extension practice, for example, to explain certain points to a farmer. Secondly, sharing the stories had a positive impact on staff morale. Thirdly, informants felt that the documentation of stories augmented the collective memory of the program. In the long term, augmenting the folk wisdom of program staff may lead to a more unified understanding of possible program outcomes. This use of stories appeared to be particularly beneficial to program staff with little field experience.

To sum up, MSC is intended to facilitate to program improvement by focusing the direction of work towards explicitly valued directions and away from less valued directions. In the case study, one year was perhaps too early to judge this, but program staff believed that the technique had led to intermediate positive outcomes. These included helping program staff and the steering committee make sense of program outcomes and understand each other's values, and drawing staff, beneficiaries and other collaborators more centrally into the evaluation process.

2. To What Extent Did MSC Contribute to Summative Evaluation?

MSC can contribute to summative evaluation by providing information about unexpected outcomes, through the selection of success stories and through values inquiry. However, we again stress that MSC is not intended to provide a stand-alone approach to summative evaluation, but can provide some very useful 'satellite' functions to complement core summative evaluation techniques

that include examination of the overall performance of the program (rather than the outer edges of experience).

Expected outcomes. MSC provides information about expected outcomes: the ‘winning’ stories provide evidence of the most successful achievements of individual clients/beneficiaries, as viewed by stakeholders involved in the process. These stories tend to be restricted to domains that correspond to program objectives rather than the open window domain (any other changes). These winning stories provide a form of performance information that can assist judgments of program merit and worth.

In the case study, funders believed the ‘winning stories’, submitted by the stakeholder committee provided valuable information about program performance. It was interesting to note that the story rated most highly by funders highlighted the type of outcomes that flavored the next round of funding. One investor commented that the stories needed more background information to fulfill his requirements and suggested that the quantifiable figures in these stories should be converted, where possible, into dollar values.

The interpretations that accompany the MSC stories also provide additional performance information. One investor suggested that these interpretations provided tacit knowledge concerning the organizational values of the program team and demonstrated considerable organizational capacity.

Unexpected outcomes. MSC is particularly good at searching for significant unexpected outcomes and then providing a process for determining the meaning of these outcomes. Judging a program usually involves determining the extent to which a program has met its predefined objectives, but should also consider unexpected outcomes. While the evaluation of many programs may benefit from examining unexpected outcomes, MSC plays a pivotal role in evaluating programs with less predictable outcomes. For example, some extension programs have deliberately

loosely defined outcomes and participatory design, often yielding a multitude of complex and diverse outcomes. These types of programs are ideally suited to evaluation techniques that involve searching for significant outcomes and then deliberating the value of these outcomes.

In the case study, stories concerning unexpected outcomes tended to emerge under the domains of ‘any other changes’ and ‘lessons learned’. These stories provided valuable information. At the round-table meeting where funders interpreted a series of stories of significant change, and during the focus group that followed, they suggested that the stories captured the complex nature of impacts as well as introducing them to a range of outcomes of which they were unaware.

Values inquiry. In the case study, funders, program staff and stakeholder committee members were often found deliberating about how to judge stories of significant change. This involved considerable dialogue about what criteria should be used to select winning stories. In the survey, the majority of respondents stated that MSC helped them to understand each other’s values. Funders also expressed surprise at the extent to which different funders used different criteria to judge the stories. Clearly, MSC helped to reveal the different values of the stakeholders involved. It is also possible that this process of making values more ‘visible’ could also have influenced the choice of success criteria used to judge the whole program.

3. To What Extent Did the Processes Encourage Stakeholder Participation?

In the case study, many stakeholders became actively involved in the MSC process. It was especially noticeable that farmer committee members engaged with (and enjoyed) the process of collecting and interpreting the stories of significant change. The funders also visibly enjoyed the process and were keen to participate in future rounds of story selection.

However, the process favored the inclusion of some stakeholders over others. Indeed, the story selection process is extremely biased in favor of those people who attended the story review sessions. In the case of the Target 10, however, the committees were usually attended by a

reasonably representative spread of farmer members, program collaborators and staff. Nevertheless, MSC is clearly embedded within the discourse of the program staff and committee members. The farmer-clients involved in the selection process cannot be considered to fully represent all farmers on matters of program practices and outcomes. MSC made no attempt to capture the opinion of farmers who chose not to participate in the program, so may not have given voice to critics of the program.

Some approaches to evaluation take steps to ensure that critics are represented. For example, the ‘fourth generation evaluation’ process proposes identifying and seeking out the views of the ‘victims’ of a program (Guba and Lincoln 1989). Including a similar process could add to the validity of MSC. We believe that staff would have been willing to discuss stories from program critics and that MSC could easily accommodate this. On several occasions, staff mentioned the need for more stories about lessons learned—or not learned—because stories about negative outcomes tended to generate a high level of discussion and learning.

Combining MSC with a process that sought out program critics could offset the bias and provide a more comprehensive evaluation outcome. Stories collected from critics could be included in the review forum, or critics could be included in the review panels.

Another element of participation relates to who is given voice within an organization. MSC employs some mechanisms for balancing unequal voices in organizations. As the process is usually embedded in a highly visible power structure, all judgments are made much more public than they might otherwise be. Those at the top of the hierarchy have to choose from menus of options created by those below them. Finally, the option of “any other changes” as a domain opens up the breadth of change that can be placed on the menu. Although the choices are never entirely free, because this happens within an organizational context, MSC gives a greater voice to those at the bottom of the organizational hierarchy than is the case with many conventional monitoring and evaluation systems.

4. To What Extent Were the Design, Data Collection and Analysis Valid?

On several occasions during the trial, people expressed doubts about the scientific validity of the outputs of the process. Some people were especially concerned with the sampling technique, the bias of the technique towards positive outcomes, and the authenticity of the stories.

Sampling technique. MSC employs purposive sampling. It provides information about exceptional circumstances, particularly successful circumstances, rather than information on the ‘average condition’ of a client/beneficiary who attends the program. While this sampling frame may not produce reliable information on which to judge the overall performance of a program, it can contribute important information to an overall evaluation. MSC aims to capture significant instances of success or failure with the purpose being to learn from these extreme stories, and ultimately to change practices to move towards success and away from failure. Thus the strategy selects those stories from which most can be learned. Patton (1990, p.170) has argued that “in many instances more can be learned from intensively studying extreme or unusual cases than can be learned from statistical depictions of what the average case is like”. In the case study program, MSC was combined with other techniques that employed a random sample to gain an understanding of the average experience of participants, as well as the extreme cases.

Bias. In the case study, MSC was biased in favor of success rather than ‘bad news’ stories. About 90 percent of stories concerned positive outcomes. However, towards the end of the trial, program staff and committee members elected to include a further domain named ‘lessons learned’, to ensure that each region would present at least one bad news story in each selection period. Several staff pointed out that there seemed to be more potential for learning from the bad news stories, and for this reason there was a move to encourage collection of these stories.

The selection process is biased towards the views of committee members and those involved in the story selection process. However, in contrast to other techniques, the bias itself is visibly

offered as another source of data about the organizational capacity of the program. The reasons behind the selection of stories are recorded and documented along with the stories themselves. The inclusion of these interpretations as another form of evaluative data affords a high level of transparency.

Authenticity. Early in the trial, participants were concerned with the authenticity of the stories, that is, ‘How do we know that they are not fictitious?’ All stories were accompanied by the names of those involved in the event, and the location of the event. This meant that it would have been possible to follow up the stories. Although such checking never occurred, there was considerable pressure by peers to record information accurately. Stories that seemed implausible, or incorrect in factual content, were not selected.

KEY STRENGTHS OF MSC

MSC can result in a multitude of positive effects on programs and be conceptualized in a number of ways. But to us, its key strength lies in the ability to facilitate a dynamic dialogue between designated stakeholders. This dialogue concerns the question “what do we really want to achieve and how will we produce more of it?”. Another strength of MSC is that people really seem to enjoy the process; this seems to be largely due to the storytelling process. It is also refreshingly different and thought-provoking. Because of these strengths, MSC is a particularly useful addition to evaluation portfolios for participatory programs that have diverse, complex outcomes, with multiple funders and stakeholders groups.

IMPROVING THE CONTRIBUTION OF MSC

To those thinking of adopting MSC, we make two recommendations. Firstly, the process itself could be improved to reduce bias and methodological weaknesses. Secondly, the technique should be used in conjunction with complementary techniques and approaches to evaluation.

Improving the process. MSC is currently being used by several international development programs and is constantly evolving as it is refined and adapted to new contexts – it is a work in progress. We offer some suggestions below for improving the process to better meet needs.

Firstly, MSC could be improved by adding a process to formally incorporate the lessons learned from the stories into long-term and short-term program planning. This could be done by asking those reporting MSC stories to offer recommendations for action derived from the stories they have selected.

Secondly, we suggest that MSC could be revised to include the voices of critics and non-participants. This could involve deliberately seeking out stories from critics of the program. Another option is to expand or modify the range of groups who select winning stories. There is no reason to restrict story selection to steering committees and investor groups. It would be possible, for example, to involve members of the public in dialogue concerning which stories they did and did not value. Conducting part of the MSC process on the Internet would enable many more people to be involved in voting for stories and explaining the different reasons behind their views.]

Thirdly, nominated stakeholders could periodically analyze the stories en masse, in addition to making judgments about the relative merit of particular groups of stories. For example, stakeholders could help to identify major themes arising from the whole spectrum of stories, including those not selected. These themes could be further analyzed, and perhaps quantified for the extent of their occurrence and compared to expected outcomes. Another possibility would be to use an approach similar to that advocated by Kibel (1999) where stories are coded against an outcome hierarchy followed by contribution analysis.

Combining MSC with other techniques. MSC has different biases to those found in conventional methods of monitoring and evaluation. For this reason, it is a particularly good tool to

use in combination with other evaluation methods chosen to offset inherent biases and meet different demands. Complementary evaluation approaches might provide:

- quantitative evidence of the spread of the emergent outcomes
- evidence of the achievement of predetermined outcomes, if these have been articulated
- evidence of the ‘average’ experience of participants, or of subgroups of participants, as well as exceptional outcomes
- information on the views of non-participants, and victims of the program.

CONCLUSION

MSC has the potential to become a useful addition to the basket of choices for evaluating programs in developed economies. It is a particularly valuable tool for improving programs with diverse stakeholder groups and values. MSC can also act as an accessory, providing a series of functions to enhance the core summative evaluation effort. In particular, MSC can help to uncover important and valued program outcomes not initially specified. This function may be especially pertinent to programs with diverse, context-specific outcomes.

The unusual methodology of MSC offers high returns. Yet the technique is still evolving and while we make some suggestions for improving the technique, there will be other useful ways to improve or adapt it to different contexts. We invite readers to join with us in exploring how MSC can be developed and creatively combined with other techniques and approaches in evaluation.

ACKNOWLEDGMENTS

We gratefully acknowledge the Victorian Department of Primary Industries, Australia. In particular we would like to thank the staff and stakeholders of the Target 10 dairy extension program and the Evaluation Support Team for their involvement and support for evaluation work.

REFERENCES

- Bennett, C. F. (1975). 'Up the hierarchy'. *Journal of Extension*, March / April, 6-12.
- Bennett, C., & Rockwell, C. (1995). *Targeting outcomes of program (TOP): An integrated approach to planning and evaluation*. Washington: United States Department of Agriculture (NOTE: 2003. The most recent information is on the website: <http://deal.unl.edu/TOP/>)
- Boje, D. M. (1991). The storytelling organization: A study of story performance in an office-supply firm, *Administrative Science Quarterly*, 36, 106-126.
- Campbell, D.T. (1969). Variation, selection and retention in socio-cultural evolution. *General Systems*, 14, 69-85.
- Dart, J. J. (2000). Stories for change: A new model of evaluation for agricultural extension projects in Australia. PhD thesis, Institute of Land and Food Resources. Melbourne: University of Melbourne.
- Dart, J.J. (1999a). 'A story approach for monitoring change in an agricultural extension project'. *Proceedings of the Association for qualitative research (AQR) international conference*. Melbourne: AQR, [online]: <http://www.latrobe.edu.au/www/aqr/offer/papers/JDart.htm>.
- Dart, J.J. (1999b). 'The tale behind the performance story approach'. *Evaluation News and Comment*, 8, no.1, 12-13.
- Davies, R. J. (1996). *An evolutionary approach to facilitating organisational learning: An experiment by the Christian Commission for Development in Bangladesh*. Swansea. UK: Centre for Development Studies, [online]: <http://www.swan.ac.uk/cds/rd/ccdb.htm>. This paper has also been published, with some variations, in Mosse, D. Farrington, J., and Rew, A. (1998) *Development as process: concepts and methods for working with complexity*. London: Routledge/ODI (pp 68-83); and in *Impact assessment and project Appraisal*, 16. No. 3, September 1998, 243-250.
- Davies, R. J. (1998). Order and diversity: Representing and assisting organisational learning in non-government aid organisations. PhD thesis. Swansea. UK: Centre for Development Studies, [online]: www.swan.ac.uk/cds/rd/thesis.htm.
- Delbecq, A. Van de Ven, A, & Gustafson., D (1975). *Group techniques for program planning: A guide to nominal group and delphi processes*. WI: Briar Press Middleton.
- Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park: Sage.
- Guijt, I., Arevalo, M., & Salsdores, K. (1998). Tracking change together. *PLA notes: Participatory Learning and Action*, 31, 28-36.
- Henry, G.T. (2002). Choosing criteria to judge program success: a values inquiry. *Evaluation*, 8, 2, 182-204.
- House, E.R., & Howe, K.R. (1999). *Values in evaluation and social research*. Thousand Oaks, CA: Sage.

- Jacobs, A. M., Fivars, G., Edwards, D. S., & Fitzpatrick, R. (1978). *Critical requirements for safe-effective nursing practice*. Kansas City, MO: American Nurses' Association.
- Kibel, B.M. (1999). *Success stories as hard data: An introduction to results mapping*. New York: Kluwer Academic/Plenum.
- McDonald, B., Rogers, P. & Kefford, B. (2003 in press) "Teaching people to fish? Building the evaluation capability of public sector organisations". *Evaluation*, forthcoming.
- Mark, M.M, Henry, G.T., & Julnes, G. (2000). *Evaluation: An integrated framework for understanding , guiding, and improving public and nonprofit policies and programs*. Jossey-bass: San Francisco.
- Nelson, R.R., & Winter, S.G. (1982). *An evolutionary theory of economic change*. Cambridge, Mass: Harvard University Press.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods*, Newbury Park, US: Sage.
- Patton, M. Q. (1997). *Utilization focused evaluation*. Thousand Oaks, US: Sage.
- Riessman, C. K. (1993). *Narrative analysis*. Newbury Park, US: Sage
- Shaw, G., Brown, R., & Bromiley, P. (1998). Strategic stories. In *Harvard Business Review*, Vol. May-June , pp. 41-50.
- Starr-Schneidkraut, N., Cooper, M.D., & Wilson S.R. (1997). Use of the Critical Incident Technique to Evaluate the Impact of MEDLINE. United States National Library of Medicine, [online]: <http://www.nlm.nih.gov/od/ope/cit.html>, [Accessed 13 November, 2000].

ⁱ The Delphi technique (Delbecq et al., 1975) is a form of interactive (postal) survey that utilises an iterative questionnaire and feedback to give participants an opportunity to revise earlier views after seeing the responses of other participants, until some desired level of consensus is reached.