The role of the evaluator in remote contexts

John Guenther

Introduction

What constitutes research? What is evaluation? While there are distinctions between the two terms, both appear to be distinct disciplines in the literature. However, both may involve similar methods of inquiry. There may be subtle differences in the purposes of evaluation and research but perhaps the differences are more related to the perceptions than the methods or the purposes. Evaluation is often seen as a process of 'assessment'—what works and why—while research is perhaps seen more as a process of 'discovery'. Some may suggest that the task of research is to be able to 'generalise' while the task of evaluation is to 'particularise'. Others may argue that evaluation begins with practice while research begins with theory. There is a sense in which these are 'chicken and egg' type arguments because indeed in each case the two notions may go hand in hand. Even within the discipline there are arguments about whether evaluation should be seen as based on constructivist or positivist paradigms (e.g. van der Knaap 2004) in much the same way that qualitative and quantitative—and mixed methods—researchers argue about the respective merits of their own methods (e.g. Burrell and Morgan 1979; Johnson et al. 2007; Morgan 2007).

Based on the foregoing, for most intents and purposes in this presentation I would argue that the term 'researcher' and 'evaluator' can be used interchangeably. Therefore we may be able to set aside what may be seen to be a semantic distinction and consider more fully the implications of their role—in this case within the context of diverse locations and situations within the Northern Territory. In this presentation I refer to myself as an evaluator because I have been contracted to 'evaluate'. However, much of my work involves.

In this presentation I want to share some of my learnings—in particular, those relating to my roles as an evaluator gathering good evidence in remote locations. These learnings come from working as part of a team, and I acknowledge them for their contribution to these learnings. In doing so I will give an overview of the kind of evaluations we have been involved with. I will also outline some distinctions in the field of evaluation research, which I believe are worth noting.

Purposes of evaluation

At one level evaluations are used for assessing program outcomes, typically using program logic models to determine: success of interventions (Patton 2002; W.K. Kellogg Foundation 2004); 'effectiveness and efficiency' (Stevens 2005); and what works and why, to inform the formative development of policy and practice (Dawe 2003). While the evaluators themselves may have a role

in designing the method of an evaluation, to a large degree the purpose of any evaluation is determined more by the commissioning organisation than by any single methodological approach (Chelimsky 2007). In the case of internal evaluations, where the purpose of evaluation may be driven by an organisation's need to improve professional practice or quality a 'community of practice' approach may be warranted (Wenger 1998). In the case of this latter purpose, the 'community' itself determines the purpose.

Evaluation approaches

At one level evaluation methods can be viewed dichotomously. That is they are either formative or summative. The summative approach can be viewed as an exercise 'to determine the overall effectiveness or impact of a programme or project' while the formative approach is designed to 'support the processes of improvement' (Clarke 1999). Proponents of 'empowerment evaluations' (e.g. Fetterman and Wandersman 2005; Fetterman and Wandersman 2007) take this a step further, arguing that such evaluations are not just participatory but can be used to 'foster improvement and self-determination'(Fetterman 2005:10). In summative evaluations, the evaluator is considered to be independent and in participatory and empowerment evaluations the evaluator will have an engaged, interactive role. However this dichotomy is not the only way of considering approaches to evaluation. Evaluators can take on the role of 'scientific expert', 'independent auditor' or 'consultative facilitator'. Patton (2000) suggests that this third role can be described in terms of 'utilization-focused evaluation' where the evaluator is the negotiator. He suggests that in this scenario 'all roles are on the table', just as methods are options. Markievicz (2005) suggests that this negotiation role is important for resolving multiple and potentially conflicting stakeholder interests. Role selection follows from and is 'dependent on intended use by intended users' (p. 430).

Stufflebeam's CIPP (Context, Input, Process and Product) Evaluation Model offers yet another way to consider different types of evaluation. Rather than trying to distinguish between types, Stufflebeam integrates formative and summative evaluation methods with an 'improvement focus', effectively doing away with the apparent dichotomy (Stufflebeam and Shinkfield 2007b). This kind of integration can however also be achieved within a framework of empowerment or participatory evaluation design. Suarez and Harper (2003) for example, show how a focus on outcomes can be merged successfully in a utilitarian approach to provide both objective assessment of a program's effectiveness and subjective and substantive improvement within a program. The evaluators in that example saw themselves as 'partners' with the program staff.

There are very good reasons for integrating methods and approaches in evaluation. Rao and Woolcock (2003) suggest that well-integrated evaluation methods, including qualitative and quantitative tools, enable generalisations to be made from findings, and 'that the strengths of one approach potentially complement the weaknesses of the other' (p. 168). Falk and Guenther (2007) contend that a rigorous qualitative or mixed methodological approach may produce findings that are no less generalisable than those which are based on quantitative methods.

The kind of evaluations I will describe in the following section are typically described as 'program evaluations'. They tend to rely on a kind of 'logic' (hence the term program logic). Program logic may be based on 'theories of change' assumptions—emphasising the theoretical foundations of a program; or an 'outcome approach' which emphasises the causal linkages between outputs and

outcomes; or an 'activities approach' which emphasises intended work plans (Patton 2002; W.K. Kellogg Foundation 2004). One of the perceived benefits of a program logic approach is that it builds an 'evidence base' (Pawson 2002). An important use of a logic model is to assist with the identification of indicators that will measure the intended outcomes. The choice of indicators is critical for determining impact. There is a tendency in some program evaluations to incorrectly ascribe process and output statements to outcomes (Mitchell 2000). There is also a risk that use of program logic may prescribe an outcomes framework that ultimately is not valid. Stufflebeam and Shinkfield (2007a) warn that evaluators using this approach may 'focus attention on theory developed early in the program and later discover that the program has evolved to be a quite different enterprise from what was theorised at the outset'.

Examples of program evaluations in the Northern Territory

The team I have been working with based at Charles Darwin University, have been engaged in a series of evaluations over a number of years. Each evaluation is based on a strategic or policy intervention. They have included:

- Australian Government funded 'Communities for Children' and 'Invest to Grow' programs (SPRC/AIFS 2005);

- A Northern Territory Government evaluation of Whole of Government Family and Domestic Violence Strategies (Learning Research Group and Department of Chief Minister 2007);

- A suite of Australian Government/Northern Territory Government evaluation of Family Violence Partnership Program (FVPP) projects

I will briefly describe this latter suite of projects. Each project is innovative and in some cases the concepts had not been employed previously in the Northern Territory. The projects are in various stages of completion and a final evaluation report is scheduled for completion in November 2008.

Peace at Home

Peace at Home is an integrated service for the Katherine/ Borroloola region with both police and community services for families dealing with family violence, including child abuse. There is a strong focus on intersectoral co-operation as well as interdepartmental collaboration, with community representatives and Non Government Organisations (NGOs) involved as partners with government.

Safe Families

Based in Alice Springs, the Safe Families project incorporates three components: Family Workers; access to unstaffed accommodation for Indigenous families fleeing violence through the operation of Family Houses; and access to staffed accommodation for Indigenous young people through the operation of a Kids' House.

Raypirri Rom

The Raypirri Rom Project is a family support program based in East Arnhem, with local workers supporting families at risk of violence using traditional mediation methods as well as linking families to services where required.

Indigenous Family Violence Offender Program (IFVOP)

IFVOP uses Indigenous facilitators in remote communities trained to deliver 50 hour programs to Domestic Violence offenders and voluntary participants. The IFVPP Project has now completed but is now proceeding with additional funding under the Northern Territory Government's *Closing the Gap* initiative.

Community Patrols

The Community Patrols project was designed to support the development/revitalisation of night/community patrols in at least eight remote communities over three years. The project has now been completed.

Interventions for Children

The Interventions for Children Project (IFCP) was designed to develop a set of resources for use in women's shelters and other sites where children accompany parents fleeing from violence, focussing particularly on those from remote communities.

Safe Houses

The Safe Houses Project has three components: building a mobile child protection team; recruiting ten Aboriginal Family and Community Workers (AFCWs), with a Top End and a Centre coordinator; and expanding the stock of safe houses available to remote communities.

Behaviour Change

The Behaviour Change Project is a program to train remote community facilitators in giving courses to groups of man and women (separately), supporting them to make positive changes in their relationships with partners, children and other family and community members. The approach is based on IFVOP but is broader-based, both in its clients and in its content.

Empowering Communities

The Empowering Communities project was designed to pilot a method to involve remote community members to identify, monitor and tackle local violence issues.

Stop It Before It Starts (SIBIS)

SIBIS is an anti-violence education program which was to be delivered in Northern Territory Government secondary schools through school counsellors to Indigenous 12-15 year olds.

The evaluation of these ten projects is based on some common threads, which apply to most of the initiatives. They are underpinned by the following operational principles.

1. The intention of each is to contribute to processes and practices that break the cycle of domestic and family violence across the Northern Territory;

2. They are built around strength-based approaches to practice;

3. There are elements of collaboration or 'Joined-up services' in each;

4. The approaches used are underpinned by 'family-focused practices', rather than singling out one group or another as victims and perpetrators;

5. They are designed to increase Indigenous and community ownership of projects;

6. They support building of a remote Indigenous workforce;

7. They work towards developing a remote employment model; and they are

8. Place-based approaches—as distinct from generic or 'blanket' approaches.

A complex mix of projects with multiple roles

This is indeed a complex mix of projects that goes beyond the breadth of a typical 'research' project. It calls for a multi-skilled team to work with multiple methodologies. The evaluations draw on several of the approaches described earlier. Some of the larger projects require several approaches, drawing on a combination of quantitative and qualitative data. These multiple roles can be difficult to play while maintaining the integrity of the data.

On the one hand the evaluator is required to act as a kind of impartial, objective observer. This role includes making external assessments using multiple stakeholder data sources (government, non-government, clients). This may seem fairly straight forward but because the role requires identification of strengths and weaknesses there is potential for the *interpretation* of the data to be contested. In many cases the issue is not the data itself but what it actually means and what does it allow you to say.

On the other hand the evaluator is required to be a 'critical friend'. This role is not objective and nor should it be. The purpose is to offer reflective practice approaches, ensuring that learnings are applied to the context. This can be a particularly useful role for program managers and officers who are wanting to critically reflect on their practice and who want to improve practice and outcomes. This role can include a capacity building function as well offering training and building evaluation tools. For example, one of my roles has been to build a database for a project so the program staff can collect data to contribute to their evidence.

Data collection issues for evaluators

I'd now like to focus on what I see as one of the major challenges for evaluators in remote contexts. This relates to gathering data—collecting evidence that is both meaningful, valid and reliable. Some of these issues may overlap with other contexts, but many are unique to work in remote areas.

1. It's not my job to collect data'

There have been several instances in the evaluations mentioned above, where the message I have received from a person—either explicitly or tacitly—involved in the project is 'it is not my job to collect data'. The rationale given for this message is that: 1)it is the evaluator's job to get the data; 2) there is no time to collect data; and 3) 'our focus is on client delivery, not evaluation'. Now of course, it may be the evaluator's role to collect data but often it can only be achieved through someone who has access to the data. In effect this is an excuse and may be a pointer to the possibility that data is not being recorded or that no person has been instructed to do this job.

2. 'You can't measure that'

A second reason for a failure to have adequate data collection systems in place is given by the blanket statement that says: 'you can't measure that'. This may reflect the attitude shown in the first issue above, which seemingly suggests that 'it is all too hard and I don't want to think about trying'. However, from my experience working in mixed methods evaluations this reason relates to two other things: 1) a failure to recognise the validity of qualitative data; and 2) a failure to identify appropriate indicators that reflect the attributable low level outcomes of a program.

3. Sometimes a recording system needs to be created

My experience working with NGOs and government agencies is that often they do not have appropriate recording systems in place to be able to capture the data that is required for either reporting or evidence base purposes. This is more likely to be about the capacity to do this kind of task than it is about a willingness to create such a system. In many cases the systems required to capture the data are not complex and a simple database or spreadsheet will suffice. If the evaluator has the required skills to do this it can be done 'on the fly' while working on other evaluation issues. However, it is important to build in training and support for the use and iterative development of these kind of tools. If the evaluator does not have the skills required, then these may have to be outsourced. One way or another, a recording system is required if evidence is to be gathered.

4. Low English language literacy and numeracy skills

For non-English speakers—particularly those in remote contexts—the level of literacy and numeracy required to understand the concepts associated with text and number based data collection tools, can be extremely daunting. While it may be possible to have data collection forms completed, questions will always be raised about the reliability of the data. The early attempts by Communities for Children to gather data from remote communities is a good example of how not to go about this. They had multiple forms, some as long as 20 pages, full of mainstream concepts, and assumptions that would have been—if someone had tried to record the data—totally misinterpreted and rendered useless. Fortunately in that instance they abandoned that method and adopted other ways of getting information they needed. I have recently been involved in a couple of projects where we are attempting to use pictogram-based tools as a way of addressing the English literacy and numeracy issue. These are still in trial phases, but they represent at least some way of creatively addressing the issue.

5. Find the right person to collect the data

What a difference it makes when you can work with a person who is keen to do the data entry work. By contrast, working with someone who dislikes this work (or doesn't see it as important) is extremely frustrating. I have observed both scenarios in government and non-government agencies alike. Some people appear to have a kind of 'block' to this kind of work and will find all sorts of reasons why it hasn't been done. On many occasions I have scheduled a visit to follow up on data collection and I will arrive—sometimes at considerable expense—to find that nothing has been done. However, when an organisation employs the right person or the right connections are made to the correct person, data collection and subsequent analysis move along very easily.

6. Government databases have their own inherent problems

Aside from the issue noted above, extracting data from government systems can be problematic. There are two reasons for this: 1) The systems are not necessarily set up to extract the combination

of variables that an evaluation sometimes requires (even though the data might be in the system); and 2) The data, once extracted may not be as reliable as would be expected. I have come across one scenario where the data I requested had to be manually accessed (i.e. physically reviewed, record by record) because there was no ability within the system to create the necessary query. In another scenario, there was no person available to be able to access the data. In another situation where I was able to access data the definitions and criteria had changed during the reporting period making longitudinal comparisons more difficult.

7. Privacy and confidentiality

Privacy and confidentiality are of course issues that must be confronted early in a project. They are ethical issues of fundamental importance. However, at times I have found that even de-identified data is not made available because of apparent privacy and confidentiality concerns. Sometimes this is related to the lack of interagency protocols and other times it is related to a fear that data may be misused. Nevertheless, these ethical issues can be a stumbling block for gathering of good evidence.

8. Organisational cultures

Having access to data sometimes depends on an organisation's culture. Some organisations are extremely protective about releasing anything to an evaluator. This may in part be due to a fear that the evaluator is going to use the data to write a 'report card' against them. Alternatively and perhaps more likely it is due to the prevailing culture within an organisation that actively discourages sharing of information with anyone outside the organisation. The contrast between these cultures is sometimes stark. One organisation will do everything it can to enable access to appropriate data (probably because they know that there is much that can be learned from the analysis of the data), while another will do everything it can do prevent access to data. Sometimes it comes down to personalities within an organisation. I have observed some organisations where one person is extremely cooperative and another is the opposite.

So how can these issues be overcome?

In many cases these obstacles are not insurmountable. I have found a few key strategies do often help.

1. Build a relationship with the people you are working with.

Building a relationship with someone takes time, but investing in a relationship during an evaluation can pay huge dividends. There are several formal and informal things I do to try to nurture that relationship:

- Take the people involved out for a cup of coffee and away from the workplace and take time to chat about what is important to them;
- Keep in regular communication by phone and email;
- Send reminders so they are not surprised by a visit and check if it is still OK to visit;
- Discuss professional issues that are outside the scope of the evaluation; and
- Share something of yourself, not just your skills.

2. Offer support to overcome the barriers.

Often the barriers to data collection are real, not excuses. What I try to do is offer whatever I can to support the people involved in an evaluation so they can overcome them.

- Suggest ways you can help, even if it costs you time and money;
- Use your expertise to the fullest—if you have skills in another area that might help, offer those skills to support the project; and
- Limit the possibility of excuses getting in the way by suggesting ways around the problems.

3. Demonstrate the value of the evaluation to the organisation.

Sometimes the barriers to data collection are related to the low value placed on evaluation by the organisation. I have observed that by helping program managers see what evaluation will do for them, they will become more cooperative and creative about how they approach data collection. Among the things I do to achieve this are:

- Explain the program logic of a program and identify the expected outcomes and show how they can be measured;
- Show that the outputs of the evaluation (literature reviews, reports, data analysis etc) can be used by the organisation to access more funding in the future;
- Work on a dissemination project together (e.g. conference presentation, journal article, seminar) where the value of what is being achieved through the project can be recognised; and
- Provide data analysis and reports that are user-friendly to the organisation and make life easier for them in other ways (e.g. for internal reporting and media releases).

4. Be patient.

Sometimes the key to getting good data is to be patient. I try to adopt the attitude that if I have done all I can then it will all fall into place (which it mostly does) or it will all fall in a hole (in which case I will hopefully be outside the hole!). The evaluator sometimes needs to find a delicate balance between gently pushing to get things moving, and patiently waiting for the right people to act.

5. Help people to think 'outside the square'.

People who are intimately involved in a project sometimes have difficulty seeing the wood for the trees. I have noted that when an evaluator comes along with a fresh set of eyes and offers their own insights it can be quite enlightening. I have on many occasions interviewed program managers who have later commented that the experience was almost 'therapeutic'—that is they were able to clearly think through the issues raised without distraction and were guided through a rich learning process.

6. Go to them, don't expect them to come to you.

People in programs are busy. This doesn't mean that evaluators are not. But if it is important to build relationships with those you are working with on an evaluation, then it is vitally important that you go and see them on their 'turf'. When you go, make sure you are on time for appointments and if you say you are going to finish by a certain time then do so. All of these strategies are designed to make the person assisting you with an evaluation, feel important—because they are.

7. Ultimately if you have a job to do, you may have to use a stick instead of a carrot.

Unfortunately there are times when the best of intentions, cups of coffee, patience, creative thinking and support do not work. If, as an evaluator you are funded to do a job and you are prevented from doing the job because someone is not doing theirs, then as a professional you may have to take measures that will not always make you popular. For example, I recently conducted a site visit to review data collection and found that the data had not been entered. I indicated as nicely as I could, that if I did not receive the data then I would have to report that I had received no data. Within a bit over a week, the data arrived in my inbox and I was able to carry on with my job and conduct some analysis—and it turned out to be a very useful set of data. While none of the following measures are my preferred way of working, there have been times when I have had to resort to:

- Discussing the matters raised with an employee's superior;
- Reporting failures in the organisation's recording systems; and
- Making adverse recommendations that relate more to systems and organisational culture than outcomes.

This list of ways of addressing data collection issues is not necessarily exhaustive but it does show that there are several strategies available to an evaluator to ensure that the data collected are complete, accurate and valid.

Conclusions

What I have tried to do in this presentation is to highlight some of the challenges and solutions related to data collection as they are associated with complex research evaluations of the kind we have been working on in diverse locations in the Northern Territory. Whether we call it research or evaluation, the issues for those engaged in discovery and learning processes in these contexts are probably much the same. The major professional risk for myself and those I am working with is that we may inadvertently 'shoot ourselves in the feet'. Another major risk is that the hard work we put in and the things we learn and find are gazumped by a changed political climate that overrides the findings we make. The report we write may then never get published or just sit on the shelf.

However, the kind of evaluations we are working on relate to practices that are innovative, cutting edge and extremely interesting from a research and evaluation perspective. In relation to the issue of gathering data, while we still have some way to go, the tools we are now developing are increasingly becoming useful as we attempt to build a sound, incontrovertible evidence base. Further, the issues we are reporting on are of critical importance for many people who have had to contend with the results of violence we are researching. Following on from this—while there are of course risks—there is a real opportunity to make a difference to the way that organisations work and ultimately deliver outcomes for their clients.

References

Burrell, G and Morgan, G, 1979, *Sociological paradigms and organisational analysis*, Heinemann, London.

Chelimsky, E, 2007, "Factors Influencing the Choice of Methods in Federal Evaluation Practice." *New Directions for Evaluation* **113** (Spring 2007): 13-33.

Clarke, A, 1999, *Evaluation Research: An Introduction to Principles, Methods and Practice* Sage Publications, Thousand Oaks.

Dawe, S, 2003, *Basing policy and practice on sound evidence*, Australasian Evaluation Society International Conference 2003, Waipuna Hotel, Mount Wellington, Auckland, Aotearoa New Zealand, 18 September 2003.

Falk, I and Guenther, J, 2007, *Generalising from Qualitative Research: Case studies from VET in Contexts*, EVOLUTION, REVOLUTION OR STATUS QUO? The new context for VET, AVETRA 2007 - 10th Annual Conference, Victoria University, Footscray Park, April 11-13, 2007.

Fetterman, D, 2005, The Heart and Soul of Empowerment Evaluation, *Empowerment Evaluation Principles in Practice*, Eds. D. Fetterman and A. Wandersman, New York, The Guilford Press**:** pp. 1-26.

Fetterman, D and Wandersman, A, Eds. 2005, *Empowerment Evaluation Principles in Practice*, The Guilford Press, New York.

Fetterman, D and Wandersman, A, 2007, "Empowerment Evaluation: Yesterday, Today, and Tomorrow." *American Journal of Evaluation* **28**(2): 179-198.

Johnson, RB Onwuegbuzie, AJ and Turner, LA, 2007, "Toward a Definition of Mixed Methods Research." *Journal of Mixed Methods Research* **1**(2): 112-133.

Learning Research Group and Department of Chief Minister, 2007, *Evaluation of the Northern Territory Government Domestic and Aboriginal Family Violence Strategies*, Department of the Chief Minister, Learning Research Group, Charles Darwin University, Darwin

Markiewicz, A, 2005, "'A balancing act': Resolving multiple stakeholder interests in program evaluation." *Evaluation Journal of Australasia* **5**(1 & 2): 13-21.

Mitchell, P, 2000, *Building capacity for life promotion: Technical report volume 1*, Evaluation of the National Youth Suicide Prevention Strategy, Australian Institute of Family Studies, Retrieved March 2006, from http://www.aifs.gov.au/ysp/yspevaluation/evalrep2chap1.pdf.

Morgan, DL, 2007, "Paradigms Lost and Pragmatism Regained: Methodological Implications of Combining Qualitative and Quantitative Methods." *Journal of Mixed Methods Research* **1**(1): 48-76.

Patton, M, 2000, Utilization-Focused Evaluation, *Evaluation Models: Viewpoints on Educational and Human Services Evaluation* Eds. Daniel L. Stufflebeam, George F. Madaus and T. Kellaghan, Springer.

Patton, M, 2002, *Qualitative research and evaluation methods*, Sage, Thousand Oaks, CA.

Pawson, R, 2002, "Evidence-based Policy: The Promise of 'Realist Synthesis'." *Evaluation* **8**(3): 340-358.

Rao, V and Woolcock, M, 2003, Integrating Qualitative and Quantitative Approaches in Program Evaluation, *Evaluating the Poverty and Distributional Impact of Economic Policies*, Eds. F. Bourguignon and L. Pereira da Silva, Washington DC, The World Bank**:** 165-190.

SPRC/AIFS, 2005, *Stronger families and communities strategy Draft national evaluation framework*, Social Policy Research Centre, Australian Institute of Family Studies, Canberra.

Stevens, U, 2005, *Bridging the Service Divide: New approaches to servicing the regions, 1996-2001*. Canberra, University of Canberra, **Doctorate of Public Administration**.

Stufflebeam, D and Shinkfield, A, 2007a, *Evaluation Theory, Models & Applications*, Jossey-Bass, San Francisco.

Stufflebeam, D and Shinkfield, A, 2007b, *Evaluation Theory, Models, and Applications*, Jossey-Bass San Francisco.

Suarez-Balcazar, Y and Harper, G, 2003, Implementing an Outcomes Model in the Participatory Evaluation of Community Initiatives, *Empowerment and Participatory Evaluation of Community Interventions: Multiple Benefits*, Eds. Y. Suarez-Balcazar and G. Harper, Binghamton, NY, The Haworth Press**:** 5-20.

van der Knaap, P, 2004, "Theory-Based Evaluation and Learning: Possibilities and Challenges."
    Evaluation **10**(1): 16-34.

W.K. Kellogg Foundation, 2004. *Logic Model Development Guide*.   Retrieved August 2005, from
    http://www.wkkf.org/Pubs/Tools/Evaluation/Pub3669.pdf.

Wenger, E, 1998, *Communities of Practice. Learning, meaning and identity*, Cambridge University
    Press, Cambridge.