Gathering Evidence: Use of Mixed Methods In Evaluation of a Program for Indigenous Clients

Jenny Williams, Hunter Valley Research Foundation

This paper discusses use of a mix of qualitative and quantitative methods in evaluation of a program that delivers community services to Aboriginal and Torres Strait Islander clients[1]. The main focus of the paper is on the use of a Computer Aided Telephone Interview (or CATI) survey with Aboriginal clients, the issues and methodological considerations that arose, and how these were dealt with in the study. The strengths and weaknesses of using the various methods with Aboriginal people and the contribution of the mix of methods to the overall evaluation are discussed.

The author acknowledges the opportunity given to the Hunter Valley Research Foundation (HVRF) to work with a community service agency to conduct the evaluation, and for their support in granting permission to present this paper. The evaluation was made possible through the contributions and cooperation of the Aboriginal staff and clients of the agency.

The detailed evaluation results and interpretation are not discussed here, as control and ownership of these are appropriately vested in the agency.

Project Brief

The brief for the project required a **process** or clarificative evaluation that focused on program processes more than outcomes. The objectives were to identify, review and evaluate all aspects of a key component of the agency's programs including work practices, relationships with internal and external stakeholders, infrastructure and resources, and organisational structure.

The brief specified inclusion of a review of current best practice in service provision, and consultation with a variety of stakeholders, including clients. To this was added – through negotiation with the agency – analysis of administrative data.

Contextual Factors

Within the broader community, Aboriginal Australians continue to suffer significant disadvantage, particularly in terms of life expectancy, and general health and welfare. This is despite attempts by governments at all levels in recent decades to specifically address the disadvantage and to make services more accessible (AIHW, 2006).

Further, whilst Aboriginal health and welfare services are now often provided by Aboriginal organisations or programs, run by Aboriginal staff for Aboriginal clients, they are still funded and administered through Western-style structures. Evaluation of such services needs to be convincing to Western-style thinking, so there is real need for an array of research methods that are acceptable to Aboriginal and non-Aboriginal communities and to decision-makers (Bullen, 2004).

There has been increasing attention to the value of using a mix of methodologies in evaluation generally. In particular, the use of a mix of **qualitative** and **quantitative** methods is seen as potentially yielding a depth of understanding of the organisation or program being evaluated as well as providing quantitative or "hard" data to decision makers (such as funding bodies).

[1] The collective term 'Aboriginal' is used throughout to refer to Indigenous Australians.

In addition, Australian Aboriginal communities have been the subject of much research, particularly in the health field. The literature reflects increasing recognition of the need to conduct research WITH rather than research ON Indigenous peoples. Participatory research has been seen as a way of making research culturally appropriate, recognising that Western "ways of knowing" and Western values are often different from those of Indigenous peoples, and also as a way of empowering the communities who are the "subject" of such research (Minore et al, 2004).

At the same time there has been explicit acknowledgement that a truly participatory research paradigm may be impractical due to factors such as time, cost, availability of sufficiently skilled Aboriginal personnel, and geography (Spooner et al, 2008). In the present study, although the service agency framed the evaluation questions and nominated some of the design elements, there were constraints on participatory research: the time-frame was tight; the agency's clients were spread across a wide geographic area, including metropolitan and rural areas (but not remote communities); and there was a need for an independent outsider to conduct the evaluation.

Lastly, whilst there has been some documentation of successful use of phone survey methodology with Aboriginal service providers, it has not been a widely used technique with Aboriginal service clients. It was felt that this aspect of the evidence gathering could present particular challenges to the evaluation, and, if successful, could have application with Aboriginal clients in other contexts.

Study Design

The final study design incorporated six components: a half-day workshop with stakeholders responsible for related components of the program; a half-day workshop with agency staff directly responsible for the program component; a "yarn" session with agency staff; in-depth phone interviews with other key stakeholders nominated by the agency; a large-scale phone survey of agency clients; and detailed analysis of administrative data.

The two workshops and in-depth interviews used a common framework designed to facilitate comparisons between the various stakeholder perspectives, and broad-ranging discussion of the program's operations and processes. The framework chosen was SWOT analysis (i.e. discussion of Strengths, Weaknesses, Opportunities and Threats).

In the workshops, the SWOT analysis took the form of a discussion among the participants in small groups. Additionally, participants brain-stormed issues they thought should be addressed in the phone surveys with clients. The discussion elicited useful advice on tailoring the phone survey, including the importance of communicating respect for respondents, and the likelihood that respondents would want to chat and ask questions of the interviewers and would perceive resistance to that as rudeness[2]. The HVRF evaluation team was also warned about the reluctance of Aboriginal people to talk on the phone, especially to people they don't know.

In-depth interviews with other key stakeholders were conducted individually by phone and included the SWOT analysis. The participants – some of them non-Aboriginal – came from a variety of different perspectives and provided a more strategic view of the program.

[2] Note that some of the issues raised would apply to survey design for any population.

The original study design was modified after the workshop with agency staff to include an informal meeting between the evaluator and frontline staff – a "yarn" session. The aim was to increase participation by the staff in the study, and make collection of the qualitative data more culturally appropriate. The session was designed as an informal "chat" to allow staff to raise and discuss issues among themselves, rather than using a structured format directed by the facilitator, and was used to augment material garnered from the structured workshops.

The HVRF team was given access to agency administrative data that tracked the program component's business activities. Sophisticated software (SPSS) was applied to "drill into" the data and report on a number of useful indicators (eg. client demographics, sources of referral, elapsed time between referral and start of service).

The Phone Survey

The evaluation team regarded the CATI survey as the most challenging aspect of the project. The HVRF has developed rigorous methodology in CATI surveys, achieving minimum response rates of 65 per cent in community surveys, and 90 per cent or more in surveys of this type where the sample is drawn from a list of known contacts. In this study, the evaluation team anticipated a substantially lower response rate.

Despite input from the workshops about what to ask and how to ask it, and an understanding of the work practices from the "yarn" session and examination of program documentation, there were three main concerns about the quality of evidence that might be gathered from this methodology:

- A structured questionnaire format is inherently at odds with Aboriginal ways of knowing and communicating
- Phone administration is at odds with Aboriginal preference for face-to-face communication
- The evaluation literature indicated that client satisfaction studies tend to produce positive results, for a variety of reasons. Further, this author's experience was that community service clients could be reluctant to make critical or negative responses partly from a concern that they might lose the service, or even jeopardise the program, if they "complained". It was expected that this would be at least as true for Aboriginal clients as for non-Aboriginal clients.

A number of strategies were adopted, in addition to incorporating the advice from stakeholders into the questionnaire design, to make the survey as culturally appropriate as possible within the bounds of methodological rigour, and to maximise the response rate so as to yield robust and reliable data. These included:

- A letter from the agency to all clients in advance, to advise that they might be phoned by an independent researcher to ask about their experiences with the service. The letter explained that they would not be identified and their answers would be confidential. It asked them to provide frank feedback, and invited them to opt out of the study.
- The involvement of workshop participants as respondents to the first pilot. The questionnaire was reworked on the basis of their feedback. A second pilot was conducted with real clients, resulting in no significant change to the questionnaire.
- The final questionnaire avoided jargon completely and used simple language and response categories; questions were framed in terms of frequency of practices rather than using scales of importance or satisfaction.
- A check was included at the start of each interview to ensure that all respondents had opportunity to know about the letter, which could be read over the phone. Confidentiality, de-identification of responses, voluntary participation and the

option to withdraw at any time during the interview were stressed to all respondents.

- Provision was made for clients who were willing to participate but unable to do a phone interview (because of language difficulty or a disability) to have a helper answer on their behalf. These data were analysed separately.

In addition, some of HVRF's standard protocols were modified and interviewers were briefed on some aspects of Aboriginal culture[3]. The standard protocols include random selection of respondents and adequate sample size to yield results with a narrow confidence interval at a 95 per cent confidence level; conduct of interviews in an on-site facility by experienced interviewers, who are trained to adopt a neutral tone of voice, follow the script exactly and politely discourage chatting with the respondent; monitoring of a percentage of interviews for quality control; piloting of questionnaires, usually two pilots, to ensure face validity and logical question flow; a minimum of six attempts to contact a selected respondent (or household), at different times and on different days (plus protocols for fax/phones, answering machines, and calls back); responses are confidential and participation is voluntary; respondents who offer comments during the interview are asked to save them until the end.

The evaluation team drew a random sample across all geographic areas serviced by the agency from a complete list of clients. The target was 300 completed interviews. A series of screening questions was included to ensure that respondents were talking about the program component under review (as community service clients can be recipients of services from several agencies simultaneously).

The standard protocols were applied in terms of sampling and sample size, contact procedures, piloting, confidentiality and voluntary participation, use of neutral tone and following the written script. The HVRF's experienced interviewers were briefed to expect that the respondents would be likely to want to chat more than non-Aboriginal respondents and that to cut this off would be likely to be seen as rudeness, resulting in a refusal to proceed. To support this, the CATI program was designed to allow relevant comments to be recorded when they were made (regardless of whether they related to the current question) rather than being deferred to the end of the interview. The interviewers were also briefed and debriefed in relation to some of the confronting social issues faced by respondents, which could arise during the interview.

Strengths and Weaknesses of Methods

The qualitative methods (workshops and in-depth interviews with key stakeholders) using SWOT analysis identified a variety of relevant issues. The same core issues were identified, from different perspectives, by stakeholder groups. This provided the evaluators with a depth of understanding of the operation and structure of the program that would not have been available either from the administrative data or from clients. The workshops also enabled participants to have a real say in the development and final form of the CATI questionnaire, without which the survey could have been asking the "wrong" sorts of questions. On the other hand, the format – particularly the use of SWOT analysis – reflected Western structure and thinking, and the evaluator remained relatively powerful in the situation, especially in relation to agency staff, so there was theoretically potential to miss some of the issues that were important to the Aboriginal participants.

[3] These were broad generalisations, which oversimplify the complexities and differences within Aboriginal culture, but represent the assumptions that underpinned some aspects of the methodology based on the advice from Aboriginal staff and stakeholders.

The in-depth interviews with stakeholders had similar value to the workshops, and the main weakness was that they were phone-based, which impacts the nature of the communication and can limit the extent of information gained.

The "yarn" session with staff was more culturally appropriate and provided much greater depth of information and understanding of the work practices and structural issues. It was also valuable, allowing the evaluator to observe interactions and relationships among the staff, and to some extent building a trust relationship between the evaluator and staff. This became an important factor in gaining access to a variety of administrative data. The main weakness of the "yarn" session was that the evaluator was non-Aboriginal and therefore lacked the subtle understanding of cultural nuances that an Aboriginal person would have had, and that the issue of relative power remained.

The analysis of administrative data yielded useful quantification of the staff's business activities and work practices, which put the issues identified in the workshops into perspective when triangulated with data from the other sources. On the other hand, the administrative data were quite limited in the kind of information provided, and an evaluation based solely on such quantitative data would have been quite unbalanced.

The main strength of the phone survey was that it did in fact yield a representative sample, and thus provided reliable quantitative data on work practices that could be triangulated with data from the other sources. Importantly, it provided data on client experiences that enabled identification of the proportions of the sample that experienced the program in particular ways; identification of the magnitude of some of the issues raised in the SWOT analyses; and the ability to extrapolate these to the client population as a whole. The comments from respondents also provided some insight into client perspectives on the program. It was not as culturally appropriate as face-to-face interviews would have been (but that would have been vastly more expensive), and both the structured nature of the questionnaire and the fact that it was done by phone limited the scope to explore issues in any depth. Also, phone administration automatically excludes clients who have no phone – however this issue did not affect the population in this study.

Three factors indicated that the phone survey results were both reliable and valid:

- The **response rate** – defined as the number of randomly selected eligible contacts who complete the survey as a proportion of those contacted who are eligible – was 96 per cent[4]. That is, only four per cent of eligible contacts refused to participate or did not complete the interview.
- The sample was a very good match for the age and sex profile of the whole client population, and the response patterns on objective measures (eg. service timing) showed a high level of "fit" with the administrative data.
- Triangulation of the CATI data with the evidence from the workshops, in-depth interviews, and analysis of administrative data yielded meaningful results.

Other factors that could have impacted reliability and validity included:

- The exclusion of clients who were confused about or could not remember the program component being evaluated. This is theoretically a potential source of bias in the results, as it is not possible to be sure that this group was not fundamentally different from those who were interviewed. For example, being

[4] Ineligible respondents were those who could not be contacted plus those contacted who were deemed unsuitable, primarily due to their inability to remember the program component being evaluated.

more confused or having memory problems could mean that they had greater difficulty negotiating "the system" than did respondents who could remember the program component, and so their experience of the program could have been different. However, there was no evidence to suggest that this was the case, and respondents whose memories were a bit hazy on some areas did not differ significantly from the rest of the sample.

- A few respondents clearly did not understand some questions, which could have impacted the validity of the results. This is always a consideration in structured phone interview – with any respondent group.
- Comments from some respondents indicated that they were still concerned that their participation could jeopardise their ability to continue to receive services from the agency.

Conclusion

Through triangulation, the evidence gathered from the various methods came together like a multi-layered jigsaw puzzle, with each source of evidence informing and increasing the team's understanding of the evidence from the others. For example, triangulation of data from the CATI survey and administrative data confirmed anecdotal evidence from the qualitative methods and indicated the magnitude of various practices.

Importantly, the success of the phone survey in gathering evidence for the evaluation suggests that this may be an additional tool in the repertoire when working with some Aboriginal client populations. That is, in a situation where true participatory research is not an option and yet it is vital to obtain evidence directly from Aboriginal clients, a rigorously administered, structured phone survey can yield valid, reliable data – provided the client population is predominantly one with phone connection. Clearly, other methods need to be used for remote communities, but there are many Aboriginal Australians who live in metropolitan or rural areas – who live in what has been described as an "inter-cultural" world (Taylor, 2003) – and who are recipients of a wide range of services that still require evaluation. It is hoped that this study will contribute to the current discourse on collaboration between Aboriginal and non-Aboriginal researchers and evaluators.

References

AIHW (Australian Institute of Health and Welfare), *International Group for Indigenous Health Measurement, Canberra 2006,* AIHW 2009.

Bullen (Hill), C. ' 'Doing' Indigenous research: reflections, questions and challenges', *Aboriginal and Islander Health Worker Journal*, vol.28, no.2, pp.9-10.

Minore, B., Boone, M., Katt, M., Kinch P. & Birch, S., 'Addressing the realities of health care in northern aboriginal communities through participatory action research', *Journal of Interprofessional Care*, vol.18, no.4, pp.360-368.

Spooner, C., Flaxman, S. & Murray, C., 'Participatory research in challenging circumstances: Lessons with a rural Aboriginal program', *Evaluation Journal of Australasia*, vol.8, no.2, pp.28-34.

Taylor, R., 'An Indigenous perspective on evaluations in the inter-cultural context: how far can one throw a Moree boomerang?' *Evaluation Journal of Australasia*, vol.3, no.2, pp.45-52.