

Evaluation capacity-building: a tale of value-adding

Diana Beere



Diana Beere is a Senior Policy Officer in the Office of the Commissioner, Queensland Police Service, and is currently on secondment to the Queensland Department of Education and the Arts. Email: <diana.beere@qed.qld.gov.au>

Evaluation capacity-building entails not only developing the expertise needed to undertake robust and useful evaluations; it also involves creating and sustaining a market for that expertise by promoting an organisational culture in which evaluation is a routine part of ‘the way we do things around here’. A challenge for evaluators is to contribute to evaluation capacity-building while also fulfilling their key responsibilities to undertake evaluations. A key strategy is to focus on both discerning value and adding value for clients/commissioners of evaluations. This paper takes as examples two related internal evaluation projects conducted for the Queensland Police Service that have added value for the client and, in doing so, have helped to promote and sustain an evaluation culture within the organisation. It describes key elements of these evaluations that contributed to evaluation capacity-building. The paper highlights the key role that evaluators themselves, especially internal evaluators, can take in evaluation capacity-building, and proposes that internal evaluators can, and should, integrate evaluation capacity-building into their routine program evaluation work.

Introduction

This paper draws connections between evaluation capacity-building and the notion of value-adding. It begins with a discussion of what is meant by the term ‘evaluation capacity-building’ before proposing four reasons why internal evaluators should actively engage in such work. In particular, internal evaluators are typically not only well placed to undertake evaluation capacity-building, but also have a responsibility to do so, as part of their responsibility to add value for their organisations. A case study of two linked evaluation projects undertaken within the Queensland Police Service will provide practical examples of how, by focusing on adding value while discerning value, internal evaluators can undertake evaluation capacity-building simultaneously with program evaluation.

What is evaluation capacity-building?

Evaluation capacity-building within an organisation is typically understood as an exercise in developing the evaluation skills and knowledge of some, or all, of the organisation’s staff, with a view to increasing their ability to undertake high-quality evaluations of the organisation’s projects and programs. However,

Compton, Baizerman and Stockdill (2002, p. 1) define evaluation capacity-building as ‘the intentional work to create and sustain continuously overall organisational processes that make quality evaluation and its uses routine’.

According to this definition, providing professional development in evaluation for staff does not on its own constitute evaluation capacity-building, although it is clearly a necessary component of it. Professional development might help to increase the quality of any evaluations undertaken, but it does nothing to ensure that evaluation skills will be used and maintained; nor to create and sustain a market for those skills; nor to ensure that the conduct and use of evaluation will become a routine organisational practice. In other words, the work of evaluation capacity-building is necessarily directed not only at the staff who are likely to be undertaking evaluations, but also at the rest of the organisation, in order to develop and maintain an evaluation culture. As Owen (2003, p. 43) has defined it, ‘an evaluation culture can be

... a vision, a more or less clear goal, and plans; also basic is purposive, attentive, and reflective work with others, in order to create and sustain ways of ensuring that evaluation studies and their uses continue to be carried out in the organization (Baizerman, Compton & Stockdill 2002, p. 105).

What is the role of evaluators in evaluation capacity-building?

For Compton et al. (2002), program evaluation and evaluation capacity-building are different, albeit complementary, ways of undertaking evaluation work. Given this, they do not perceive evaluation capacity-building to be part of the everyday work of either internal or external evaluators, and do not expect all program evaluators to undertake such work. Certainly, evaluation capacity-building should not be the responsibility of evaluators alone—indeed, Baizerman et al. (2002) suggest that given their typically limited authority and

power, evaluators need allies and collaborators to undertake such work effectively. However, evaluators clearly have a key role to play, and are key stakeholders in evaluation capacity-building.

Moreover, it is a central argument of this paper that

internal evaluators, in particular, will and should make evaluation capacity-building an integral part of their everyday evaluation work. The case study that follows can help to explain how internal evaluators can do this, but first it is necessary to explain why they should do so.

Why should internal evaluators engage in evaluation capacity-building?

One reason why internal evaluators should actively participate in evaluation capacity-building is that it is clearly in their own best interests to direct at least some of their efforts routinely to fostering and sustaining an organisational culture in which evaluation information is valued, evaluations are sought, and evaluation findings and recommendations are used. Without a continuing market for evaluation within their own organisations, the ongoing employment of internal evaluators is at risk, regardless of the quality of their work and the strength of their evaluation skills and experience. In this regard, internal evaluators should heed the point made by Stockdill et al. (2002) that evaluation capacity-building is a continuous project, one that can never be said to be complete. Evaluation units might be established, but they can always be shut down again, and the employment

... good program evaluations can help to create a market for future evaluation projects, and can thus support evaluation capacity-building efforts.

regarded as a commitment to roles for evaluation in decision-making within an organisation’.

According to Stockdill, Baizerman and Compton (2002), program evaluation and evaluation capacity-building are not the same thing. Indeed the actual practice of program evaluation can sometimes undermine evaluation capacity-building efforts. For example, evaluations that are poorly conceptualised and conducted, have little utility, focus inappropriately on accountability, and/or lead to program closures and job losses can result in perceptions of evaluation being a negative process and alienate potential evaluation clients (Stockdill et al. 2002; Weiss 1999).

On the other hand, good-quality evaluations can help to overcome negative perceptions about evaluation by demonstrating the range of ways in which evaluation can contribute to program and organisational improvement. By fostering an appreciation of evaluation, good program evaluations can help to create a market for future evaluation projects, and can thus support evaluation capacity-building efforts.

However, evaluation capacity-building itself is, according to the definition given above, necessarily intentional work, not just a coincidence of good fortune or hard work. Basic to the intentional nature of evaluation capacity-building are:

of evaluation staff terminated. This is arguably a real risk if internal evaluators leave the work of evaluation capacity-building to others, rather than actively participating in it themselves.

Even if internal evaluation units are not shut down, at least some participation by those units in evaluation capacity-building is likely to be necessary to justify their requests for professional development for staff, and for additional human and/or other resources. It seems that internal evaluators, particularly in the public sector, are often faced with shortages of staff, skills and other resources. They are not alone in this, but it must be recognised that so long as there is no strong internal demand for high-quality evaluations, even those organisations that have taken the step to establish internal evaluation units are unlikely to see any value in allocating additional resources to them.

It should be noted that if the above arguments for internal evaluators engaging in evaluation capacity-building seem self-serving, such engagement can also be seen as demonstrating their commitment to, and belief in the importance of, the practice and use of evaluation. Indeed, provided that they always act ethically, the engagement of evaluators (whether internal or external) in capacity-building efforts is strongly consistent with the overarching aim of the Australasian Evaluation Society (AES) (and the aims of similar associations of evaluators overseas) to improve the theory, practice and use of evaluation. So there is a second reason why internal evaluators should engage in evaluation capacity-building, particularly if they are members of the AES or similar professional bodies.

A third reason why internal evaluators should undertake evaluation capacity-building is that they are typically best placed to do so. In an analysis of four case studies of evaluation capacity-building, Baizerman et al. (2002) found that participation in such work is much easier for internal evaluators than for other evaluation practitioners. This should not be surprising, given that internal evaluators are much better placed to act, on a more or less daily basis, as advocates for evaluation within their organisation; to offer an evaluation perspective on the organisation's programs and activities by participating in internal committees and working groups; and to provide consultancy and/or training services in relation to evaluation issues, approaches and methods. These are all key aspects of evaluation capacity-building (Stockdill et al. 2002).

Moreover, internal evaluators typically have an advantage over external evaluators in terms of knowledge of their organisation and its culture and practices. Such knowledge is essential to effective evaluation capacity-building, because—like program evaluation—such work is highly context-dependent. Just as good-quality evaluations are tailored

according to the nature of the program and the needs and purposes of the clients and stakeholders, evaluation capacity-building efforts must take account of organisational contexts, including structure, culture, roles and skills. A 'one-size-fits-all' approach to evaluation capacity-building is inappropriate (Stockdill et al. 2002).

The fourth reason why internal evaluators should actively participate in evaluation capacity-building—and the reason why they arguably have a responsibility to do so—relates to their responsibility, as internal evaluators, to seek constantly to add value for their organisations. In case this responsibility is not self-evident, it follows from the premise that organisations establish internal evaluation units because they are interested in improving project and program outcomes, and hence organisational outcomes. In other words, internal evaluation units exist because of their perceived potential to add value for organisations. Thus, whether or not this is made explicit by the organisation, internal evaluation units arguably have a responsibility to realise that potential to the greatest extent possible as they undertake their program evaluation work. This means attending not only to discerning the value of projects and programs, but also to adding value by identifying ways to improve program and/or organisational outcomes. It also means engaging in evaluation capacity-building, which adds value by making the practice and use of evaluation routine aspects of 'the way we do things around here'. This in turn enhances the ability of organisations to realise their goals.

Can internal evaluators feasibly undertake evaluation capacity-building?

There are at least four reasons, then, why internal evaluators should actively participate in evaluation capacity-building: because they have a stake in the outcomes; because it is consistent with their commitment to the professional standards of the AES (or similar evaluation associations); because they are well placed to undertake evaluation capacity-building; and—most importantly—because they have a responsibility to their organisations to do so.

But how feasible is it, in practice, for internal evaluators to engage in evaluation capacity-building? Compton et al. (2002) see such work as different from, albeit complementary to, the program evaluation work that internal evaluators are typically mandated to undertake. Moreover, as noted above, internal evaluators, particularly in the public sector, often have limited staff, skills and resources, and must struggle even to complete good-quality program evaluations, without engaging in evaluation capacity-building as well.

However, Compton et al. (2002) suggest that internal evaluators can feasibly be practitioners of both evaluation and evaluation capacity-building, by focusing not only on responding to requests for evaluations, but also on how their evaluation work can contribute to developing and sustaining an evaluation culture within their organisations. These simultaneous undertakings require an orientation both 'to the specific evaluation and to how this specific program evaluation contributes to the [evaluation capacity-building] process within the larger organization and its policies, programs and practices' (Stockdill et al. 2002, p. 11).

Thus, for internal program evaluators, evaluation capacity-building remains distinct from program evaluation in the sense that it is an intentional addition to the fundamental work of discerning value. It need not be distinct in the sense of being separate, or even necessarily separable from the work and therefore does not necessarily require additional resources. Rather, what is most important is a focus on adding value.

How can internal evaluators realistically engage in evaluation capacity-building?

It is argued above that evaluation capacity-building is an undertaking that adds value to organisations. In this section, it is proposed that internal evaluators can integrate evaluation capacity-building into their program evaluation work by focusing consistently on adding value while discerning value.

To explain this, it is useful to refer to Weiss's (1999) discussion of evaluation use, and particularly her observation that evaluators, in expecting their work to make a difference, often puzzle over why their findings and recommendations are apparently ignored. She argues that evaluators need to be realistic about what their work can achieve, and also to recognise that such work frequently does make a difference over time, by contributing to what she calls 'enlightenment' and, thereby, having a cumulative, long-term effect on policy-making.

This should not be understood as a recommendation that evaluators just continue with basically the same ways of undertaking program evaluation work, perhaps with some enhancements in terms of methods. It is important to understand that being realistic about the possible effects of their work does not preclude evaluators from looking for realistic ways to achieve more, both for themselves—in terms of the satisfaction that comes from having made a difference—and for their clients, whether internal or external.

For example, if it is clear that, for political or other reasons, a program is not likely to be abandoned, regardless of the findings of an outcomes evaluation, then evaluators might instead undertake a

formative evaluation, with a view to identifying ways of improving program outcomes. This might mean challenging the client's expectation of an outcomes evaluation, but in cases where such an evaluation is inappropriate or likely to have little utility or value, this is arguably the professional responsibility of evaluators—particularly, as suggested above, internal evaluators, who should be committed to adding value for their organisations by maximising the utility of their evaluation work. As Weiss (1999, p. 471) has pointed out, 'utility is the primary purpose of evaluation; it is the name of the game'.

Granted, external evaluators may not be in a position to persuade clients whose minds are set on an outcomes evaluation to agree to such a change of approach—although they should at least attempt to do so. However, internal evaluators should be well placed to negotiate alternative approaches with their clients, particularly if they emphasise how a formative evaluation can add value for both the program and the organisation. This kind of approach from the outset of negotiations for an evaluation enables evaluators to educate clients about the various possible purposes of program evaluation, the different forms and approaches associated with different evaluation purposes (Owen with Rogers 1999), and the different kinds of information that good-quality program evaluations can provide. It is in most cases a mistake to assume clients understand these things; typically they do not, or they have only basic understandings of evaluation. Thus it is important that evaluators engage in this kind of educative work—which in itself contributes to capacity-building—every time they undertake a program evaluation. If evaluators do not take every opportunity to educate clients and potential clients about evaluation, it is difficult to see who will.

A case study

This final section of the paper presents a case study of evaluation capacity-building within the Queensland Police Service (QPS). The case study demonstrates how, in undertaking evaluations of the QPS Shopfront Police Beat and Neighbourhood Police Beat programs, the Service's internal evaluation unit engaged in evaluation capacity-building by focusing on adding value for the internal client and for the QPS as a whole while discerning value.

The Shopfront and Neighbourhood Police Beat programs represent two of the main ways in which the QPS has implemented the concept of beat policing in Queensland. A shopfront beat, as the name suggests, is a small police establishment that operates from a shop located in a shopping complex or central business district. It is responsible for dealing with policing problems within the

defined beat area—its main clients being the owners, managers and staff of commercial establishments in the area, as well as shoppers and other visitors to that area. A neighbourhood beat is a similarly small establishment with similar responsibilities for dealing with local policing problems. However, in this case the beat area is a defined residential area, and the beat officer operates from an office attached to the house in which he or she lives. In both cases, beat officers are expected to become part of the beat community; to take ownership of the beat area and its policing problems; and to work in partnership with the community to develop and implement policing strategies that proactively address local crime and disorder problems.

The two programs have similar but slightly different objectives. Briefly summarised, the aims of the Shopfront Police Beat program are to:

- enable easier communication between the community and police
- reduce fear of crime within the beat area
- increase the beat community's sense of personal safety
- increase the perceived risk of apprehension for committing an offence within the beat area.

Similarly, the aims of the Neighbourhood Police Beat program are to:

- increase police–community communication
- encourage problem-oriented and partnership approaches to chronic local policing and community problems
- reduce crime and disorder problems within the beat area
- increase the beat community's sense of personal safety
- increase community satisfaction with the police.

The broad brief from the clients—the program managers—was to evaluate both programs, the expectation being that outcomes evaluations would be conducted. However, discussions between the internal evaluators and the clients revealed that outcomes evaluations were unlikely to have any value or utility for the QPS. Both programs had strong political and public support, and in this context neither was likely to be abandoned, even if the evaluation findings were to reveal that their expected benefits were not being fully realised. Yet it was clear that the program managers had some concerns about the extent to which the programs were achieving their objectives, and were keen to ensure that they operated as effectively as possible. Notwithstanding the considerable volume of international literature on beat policing, there were few useful guidelines available to support the

effective operation of shopfront and neighbourhood police beats in the Queensland context.

In this case it was quite easy for the evaluation team to convince the clients that outcomes evaluations would be of little value, because there was an additional practical impediment to focusing on outcomes. A lack of initial consideration of the need for evaluation in the program planning and implementation stages meant that, for all but one of the objectives of the two programs, there were no clear and reliable data available to demonstrate achievements. Moreover, while it would have been technically possible to measure performance against the objective of reducing the incidence of reported crime in neighbourhood beat areas, it was not feasible to do so given the available resources. At the time, measurement of achievement of this objective would have involved the labour-intensive manual extraction of several large sets of crime statistics, based on offence addresses. While crime statistics are routinely collected for all of Queensland, they are only disaggregated to the level of police divisions, and are not readily available for beat areas, which are relatively small areas within larger divisions (data analysis tools have since been developed to address this problem).

The challenge for the evaluation team, then, was to formulate plans for evaluations that: could feasibly be undertaken with the available resources; would meet the program managers' needs for information to enhance the effectiveness of the two programs; and would have maximum utility and value to both the clients and the QPS as a whole. To meet that challenge, the evaluation team worked closely with the program managers to develop plans for two process evaluations. These evaluations sought to describe the characteristics of 'model' shopfront and neighbourhood beats, and to identify barriers to the optimal operation of each type of beat. The key methodological elements in each case were a review of literature on beat policing, in-depth on-site interviews at a sample of beats and associated police stations, and observations of salient environmental features of the selected beats.

As suggested earlier in this paper, the process of working with the clients to develop evaluation plans that were better suited to the context and needs of the two programs was in itself one that contributed to evaluation capacity-building. In this case, the process helped to increase the clients' understandings of the possible benefits and uses of evaluation, as well as the kinds of information needed to conduct particular types of evaluations, and the benefits of planning for evaluation from the outset—rather than several years after program development and implementation.

However, the evaluations also helped to build an evaluation culture among another significant

stakeholder group. Stockdill et al. (2002) cite research findings that evaluation capacity-building efforts are sometimes undermined by too great a focus on the needs and interests of high-level stakeholders, such as funding bodies and program managers, with those of other stakeholder groups marginalised. In the case of the Shopfront and Neighbourhood Police Beat evaluations, the evaluation team recognised that capacity-building efforts needed to focus not just on enhancing the program managers' appreciation of the benefits and uses of evaluation, but also on increasing the orientation of beat officers themselves to evaluation. In particular, given the expectation that beat officers will adopt problem-oriented approaches to policing problems in their local communities, evaluation capacity-building efforts needed to be directed towards developing beat officers' abilities to undertake evaluation work of their own, for the purpose of assessing the effectiveness of their problem-solving strategies.

The two evaluations contributed in several ways to evaluation capacity-building in this regard. First, both evaluation reports were presented in the form of guides to best practice, so that, rather than gathering dust on shelves, they could serve as practical handbooks for beat officers and their supervisors, as well as for the program managers. Both guides emphasised the importance of evaluation as the key to ensuring best practice in beat policing, noting that it is an essential component of the SARA (Scan—Analysis—Response—Assessment) problem-solving model promoted across the QPS to support the philosophy of problem-oriented policing. In this respect the internal evaluators were able to utilise their knowledge of the organisational and cultural context of the QPS, linking the emphasis on evaluation with a problem-solving model and a policing philosophy that were already familiar to beat officers and are strongly supported by QPS senior management.

In addition, the guide to best practice in shopfront beat policing (which was completed about 12 months ahead of its companion guide) became a major resource in the development of a one-week course to address the training needs of beat officers. These needs—identified in the evaluations—included basic training in evaluation, so the evaluation team argued successfully for the inclusion of such training in the beat officers' course.

The evaluation training component was developed and, for the first few beat officers' courses, delivered by members of the internal evaluation unit (subsequently, as program staff developed their understandings and skills in evaluation, they took over delivery of this component of the training). The training incorporates a

practical exercise in developing an evaluation plan, which follows earlier practice in applying the Scan and Analysis components of the SARA problem-solving model to a simulated policing problem. To emphasise the link between project planning and evaluation planning, the evaluation planning exercise is undertaken in conjunction with planning an appropriate set of policing responses to the simulated problem. The exercise also introduces beat officers to a one-page worksheet developed by the evaluation unit as a simple framework to assist individual beat officers (and potentially other police officers) in planning evaluations.

The evaluation worksheet, together with a completed example, was subsequently included in the best practice guide to neighbourhood beat policing. The sheet has a dual purpose, thus making an additional contribution to evaluation capacity-building. Its use by beat officers will enable evaluation information about local beat policing projects to be collated and summarised, thus providing a useful source of data for future evaluations of either individual police beats or the beat programs as a whole.

The evaluation training, with refinements over time, has now been delivered to several cohorts of beat officers, and further courses are planned. In delivering the training, evaluation unit and program staff have promoted the use of both the best practice guides and the evaluation planning sheet, which have also been made available to other police officers statewide, via the QPS intranet. The evaluation unit has also been marketed as a source of advice and assistance to beat officers undertaking their own project evaluations. Finally, participants receive a comprehensive course manual as a practical resource to guide their everyday beat policing work, and this too includes basic information and resources to aid them in evaluating the policing strategies and projects they implement in their local areas. Again, this information was prepared by the internal evaluation unit.

The evaluation unit undertook all of these evaluation capacity-building efforts in close collaboration with the evaluation clients, the program managers. An additional outcome has been a continuation, well beyond the life of the two evaluation projects, of the close relationship between the two areas that developed during those projects. This relationship is now so well established that it has survived several changes of staff, and this has proved to be important in sustaining the evaluation capacity-building efforts that began with the development of the two evaluation plans. As noted earlier in this paper, while evaluation capacity-building can be undertaken by internal evaluation units, they can rarely achieve much on their own. Moreover, it can never be assumed

that the work of evaluation capacity-building is complete, so collaborations and alliances to support this work must be sustained.

Concluding comments

The evaluation capacity-building strategies outlined in the above case study were developed in the context of internal evaluations of specific programs operating in a specific organisational context—one that, as suggested above, is already supportive of evaluation capacity-building. A ‘one-size-fits-all’ approach is no more appropriate for evaluation capacity-building than it is for program evaluations, so while the case study might help to generate some ideas, the strategies outlined here would not necessarily be useful or effective in other organisational contexts.

The key message to be drawn from the case study, and from this paper overall, is that internal evaluators can integrate evaluation capacity-building with program evaluation feasibly and effectively, and have a responsibility to do so, consistent with their responsibility to add value for their organisations. A practical approach to this challenge is to maintain consistent dual foci on adding value to improve programs and projects while at the same time discerning their value.

References

- Baizerman, M, Compton, DW & Stockdill, SH 2002, ‘Summary and analysis of the case studies: Themes across the cases’, *The art, craft, and science of evaluation capacity-building: new directions for evaluation no. 93* (pp. 101–107), Jossey-Bass, San Francisco.
- Compton, DW, Baizerman, M & Stockdill, SH (eds) 2002, *The art, craft, and science of evaluation capacity-building: new directions for evaluation no. 93*, Jossey-Bass, San Francisco.
- Owen, JM 2003, ‘Evaluation culture: a definition and analysis of its development within organisations’, *Evaluation Journal of Australasia*, vol. 3, no. 1, pp. 43–47.
- Owen J M with Rogers, P 1999, *Program evaluation: forms and approaches*, 2nd edn, Sage, London.
- Stockdill, SH, Baizerman, M & Compton, DW 2002, ‘Toward a definition of the ECB process: a conversation with the ECB literature’, *The art, craft, and science of evaluation capacity-building: new directions for evaluation no. 93* (pp. 7–25), Jossey-Bass, San Francisco.
- Weiss, CH 1999, ‘The interface between evaluation and public policy’, *Evaluation*, vol. 5, no. 4, pp. 468–486.