

A fundamental choice: internal or external evaluation?

Melissa Conley-Tyler

Melissa Conley-Tyler is an Associate with the Centre for Program Evaluation at the University of Melbourne. She became involved in evaluation after a decade of working in the non-profit sector in the US, South Africa and Australia, primarily in conflict resolution and violence reduction.

While working on this paper Melissa has conducted five evaluations as an external evaluator as well as two internal evaluations. She has a Postgraduate Diploma in Evaluation as well as degrees in Law, Arts and a Masters in Law and Diplomacy from the Fletcher School.

An organisation wishes to evaluate one of its programs. It can ask a staff member or hire someone outside the organisation. Which should it choose?

Surprisingly little guidance is available for this common scenario. A review of 30 texts dealing with organisational performance and evaluation shows that too often the issue is assumed one way or the other.

Management texts aimed at business and organisational audiences tend to presume that evaluation is conducted by internal evaluators, usually managers. By contrast the specialist evaluation literature almost always proceeds from the opposite assumption: that evaluation is undertaken by external evaluators.

This paper proposes a series of measures for comparing the strengths and weaknesses of internal and external evaluators. These include cost, knowledge, flexibility, objectivity, accountability, willingness to criticise, ethics and utilisation of results.

A set of guidelines is offered to assist organisations in choosing between internal and external evaluation in each particular case.

A common question faced by organisations wishing to evaluate their programs is whether the evaluation should be undertaken by staff members ('internal evaluators') or non-staff members ('external evaluators').¹

Given how common this scenario is, it is surprising that the choice between internal and external evaluators has not been the subject of much critical debate. Too often the issue is assumed either one way or the other without discussion of the issues involved.

This paper deals with two main strands of literature which can broadly be divided into 'management literature' (aimed at business and organisational audiences) and 'evaluation literature' (aimed at professional evaluators). A conceptual framework is offered that explains the apparent division in assumptions of these two literatures.

Finally, the positives and negatives of internal and external evaluators are discussed in detail in order to draw a set of guidelines that can be used when deciding the best option for a particular case.

Literature review

Management literature

A wide literature focuses on evaluation as a means of improving organisational performance in the business, government and the community sector. The aim is to confirm a place for evaluation as an 'integral aspect' or 'essential part' of performance improvement (Love 1991, p. 1).

Often such works offer guidance on evaluation planning, data collection and reporting mechanisms as 'an indispensable tool' for managers concerned with performance improvement (Love 1991, p. 1). A range of specific planning and evaluation tools are offered by Globerson, Globerson and Frampton (1991), Love (1991), Goldberg and Sifonis (1994), Sluyter (1998), and Czarnecki (1999).

Generally these works assume that evaluation for organisational improvement is conducted by internal evaluators, more specifically, by managers. There is rarely any discussion of when, or how, external evaluators should be used.

A good example is Globerson, et al. (1991). Defining middle management as an 'endangered

Generally these works assume that evaluation for organisational improvement is conducted by internal evaluators, more specifically, by managers. There is rarely any discussion of when, or how, external evaluators should be used.

species', Globerson et al. offer a new role for middle managers as internal planners and evaluators who can exercise organisational control and improve performance: 'These managers should be able to evaluate performance by using specific performance criteria, measuring actual results, and comparing the results with expectations' (Globerson et al. 1991, p. ix). The elaborate structure of planning, setting standards, measurement and comparison offered by the authors rests on the unstated assumption that only management is involved (p. 15).

Similar works include Love (1991), Sluyter (1998) and Czarnecki (1999) who generally assume that managers rely on internal evaluators for measurement.

Sluyter (1998) presumes that internal evaluators should measure 'indicators of quality and stability' for performance improvement (pp. 68–70) while Goldberg and Sifonis (1994) presume that monitoring and evaluation processes such as surveys, activity-based costing, trend watching and benchmarking will be carried out by internal evaluators.

Love (1991) presents a brief discussion of the advantages and disadvantages of internal evaluation (pp. 4–5), conceding that an alternative is at least possible, while Czarnecki (1999) suggests (once) that a consultant 'with a strong statistical background' may be helpful to an internal evaluator in analysing customer satisfaction surveys (p. 130). No detailed discussion of the

relative benefits of internal and external evaluators is attempted. Unusually, Goldberg and Sifonis (1994) discuss whether there might be some benefit in 'bringing someone in from outside', at least for the planning process, and offer two circumstances in which this might be advisable: lack of staff expertise and 'turf battles' (p. 27).

Generally, however, the use of an internal evaluator is presumed in the management literature surveyed.

Evaluation literature

By contrast, the literature focused at professional evaluators almost always proceed from the opposite assumption that evaluation should be undertaken by external evaluators.

Standard textbooks such as Rossi, Freeman and Lipsey (1999, pp. 54–61) and Posavac and Carey (1997, pp. 89–92) offer much good advice to evaluators on setting up good relationships with clients and reducing conflict with stakeholders, silently presuming that the evaluator is external to the organisation being evaluated.

Cronbach et al. (1981) presume that the evaluator is separate from the 'client' (pp. 17–18), although they concede that program evaluations 'vary in the intimacy of the evaluator's relationship' with the client.

A number of other authors also use this standard division including Perloff (1979), Gunn (1987), Nowakowski (1987) (mentioning shared internal and external evaluator partnerships as something 'we normally do not see in the conventional literature', p. 4), Meyers (1981) and Scriven (1997) (cf. discussion on pp. 71–2 and p. 487 outlined below).

A good example of these assumptions at work is the classic book by Rossi and Freeman (1982) and a later edition, Rossi et al. (1999, pp. 33–4). In discussion of the question of 'Who can do evaluations?', the authors discuss the role and background of evaluation experts, noting that 'as in any other professional field, evaluators have developed their own vocabulary' (Rossi & Freeman 1982, p. 50).

The aim of the 1982 book is perhaps most revealing:

It is the purpose of this book to provide an introduction to the field ... to learn how evaluations are conducted. It is but a start along the pathway to becoming a technical expert in evaluation. Our aim is to provide persons faced with the administration and management of human resource programs with sufficient understanding of evaluation tasks and activities to be able to judge for themselves what kinds of evaluations are appropriate to their programs and projects, and to comprehend the results of completed studies relevant to their organisation. (p. 50)

At times, the authors sound like representatives of a high priesthood of evaluators—involving a select group, rather than presenting skills that can

be used for all. The only concession is the need to educate 'clients' to comprehend the results that 'experts' provide them with—thus raising the value of the evaluation profession.

In a similar vein, Nowakowski (1987) stresses the importance of 'educating more and better consumers who value and can use evaluation services' (p. 2) in introducing a collection of essays on client perspectives on evaluation. This will have the effect of positioning the evaluation field 'to influence the market for evaluation services' (p. 2). With this aim, it is not important to compare the suitability of internal and external evaluators: administrators apparently need to know how and when to hire evaluators/commission an evaluation, but not how and when to undertake evaluation themselves.

Even works such as Rothwell and Cookson (1997), who note that their principles of program planning are equally applicable to internal and external evaluators, do not review the positives and negatives of selecting between 'employees' or 'contractors' (p. 5). The discussion of 'determining program success' through evaluation presumes that the evaluation will be conducted by the program planner who designed the program (pp. 165–93). There is no discussion of when evaluation by someone other than the program planner might be useful.

A conceptual framework: audiences and evaluators

In order to understand this divergence between the management and evaluation literatures, some explanation is needed. Owen with Rogers (1999) provide a useful framework to conceptualise the choice between internal and external evaluators and the reasons why some theorists would favour one choice over others (pp. 135–6).

They identify four types of resource arrangements for an evaluation depending on the nature of the evaluator (outsider or insider) and the nature of the audience for the evaluation (outsider or insider):

- insiders for insiders
- insiders for outsiders
- outsiders for insiders
- outsiders for outsiders.

Within this framework, it is easy to see why the literature discussed above mostly falls within one or another of these four configurations.

As outlined earlier, most management literature presumes that evaluations are produced by insiders for insiders and thus focuses almost entirely on the role of internal evaluators. While some corporate evaluations are produced by insiders for outsiders (such as annual reports) or outsiders for outsiders (such as audited accounts), these are not discussed in the literature reviewed above.

For example, it is not surprising that Czarnecki (1999) focuses on internal evaluators since he

sees two main uses of evaluation for his business audience: 'to support corporate improvement efforts and as part of a performance appraisal process' (p. 14). Both of these are internal measurements that are reported to internal audiences (company managers), although they will impact eventually on the way the company is perceived by external actors.

By contrast, the evaluation literature surveyed above presumes that evaluations are produced by outsiders for an audience of either outsiders or, at best officials and the 'intelligent layperson' (Rossi et al. 1999, p. ix). Not coincidentally, the most likely readers for such books are external evaluators.

The only problem with the division in the literature based on these assumptions is that it does not assist in deciding between the four different arrangements. No guidance is offered by most texts to organisations that are seeking to choose between these models.

Thankfully, although it is rare to see the rival merits of these four resource arrangements specifically discussed, there has been some treatment of the issue. The most useful discussion of these questions is in Meyers (1981), Braskamp, Brandenburg and Ory (1987), Weiss (1998), Love (1991), Mathison (1994), Newman and Brown (1996) and Owen with Rogers (1999).

These authors are used to compare the positives and negatives of internal and external evaluators in the following section of the article.

Strengths and weaknesses of internal and external evaluators

Using the framework outlined above, it is possible to identify the key choice to be between evaluation by 'outsiders for outsiders' or 'insiders for outsiders.'

Factors that might influence the choice between these evaluation models are:

- cost
- availability
- knowledge of program and operations
- knowledge of context
- ability to collect information
- flexibility
- specialist skills and expertise
- objectivity
- perceived objectivity
- accountability for use of government funds
- willingness to criticise
- utilisation of evaluation
- dissemination of results
- ethical issues
- organisational investment.

Since a focus on these factors involves value judgements, it is important at this point to note the difficulty in formulating a standard of value of evaluation. Setting up a test for the validity of evaluations is a complex theoretical task.

For example, House established three criteria for validity of an evaluation: whether it is true, credible and normatively correct (House 1980, p. 255). He asserts that there is no one right approach to evaluation—the closer the evaluation approach to a particular situation, the more likely it is that it will be accepted. By contrast, Kushner (2000) values evaluations according to the ‘reality’ they create for the less powerful while Patton (1997, p. 21) privileges ‘real and specific’ utilisation of the evaluation.

The definition that this paper will use in judging the benefits of internal and external evaluation is that used by Cronbach et al. whereby:

... excellence ought to be judged against how evaluation could serve a society ... An evaluation pays off to the extent that it offers ideas pertinent to pending actions and people think more clearly as a result. (1981, p. 65)

This requires particular attention be paid to the views of the audience for the evaluation concerning the relevant strengths and weaknesses of the two types of evaluators.

Cost

Some authors argue that internal evaluators have an advantage over external evaluators in terms of cost (summarised in Cummings et al. 1988, pp. 60–1). By contrast, Mathison (1994, p. 301) asserts that careful use of an external evaluator can be a more cost-effective option than an internal evaluator. Both are correct.

Hour for hour, the cost of an internal evaluator is usually less than for a consultant. However, it can be very expensive to maintain idle evaluation capacity if evaluators are not needed and cannot undertake other projects (Weiss 1972, p. 35).

The difference in cost will need to be estimated in each case.

Availability

Another advantage of internal evaluators over external ones is that they are likely to be readily available: they are ‘there, ready and able to take on the job whenever evaluation is wanted’ (Weiss 1972, p. 35).

However, at times the same may be true of external evaluators, especially where an organisation possesses existing relationships with evaluators or has established an evaluation panel.

The best method of sourcing a suitable evaluator will need to be assessed in each case.

Knowledge of program and operations

Most evaluators stress the importance of gathering detailed knowledge about the program being evaluated in order to ensure that the evaluation reflects the program and its context. Posavac and Carey note that stakeholders are:

... far more likely to give recommendations a respectful hearing if they feel the evaluation team understood the population served, the agency housing the program, the program itself, and the staff. Such understanding cannot be gained by brief meetings. (1997, pp. 251–2)

This means that internal evaluators have an advantage over external evaluators since they work in the environment in which the program operates and may have been personally involved in some part of the program planning (Weiss 1972, p. 38). As part of the organisation, internal evaluators have firsthand knowledge of the organisation’s philosophy, policies, procedures, personnel and management. This allows selection of evaluation methods ‘tailored to the reality of the organization’ (Love 1991, p. 4).

However, this advantage can often be just a time advantage: external evaluators should also be able to obtain knowledge about an organisation’s program and operations given sufficient time. Only in the most complex of environments might it be impossible for an external evaluator to gain the necessary knowledge. However, the time involved may mean that cost is a factor as the organisation involved will be paying for the evaluator to gain ‘background’ that it already has.

Knowledge of context

As well as detailed knowledge about the program being evaluated, it may also be useful in some cases for the evaluator to have an understanding of cultural and political factors that affect the operating environment of the organisation being evaluated (Braskamp et al. 1987, p. 66).

In the case of a university, this might be sensitivity to university politics (Shapiro & Blackwell 1987, p. 61) while in the case of a religious institution it might be the need to be ‘attuned to the uniqueness’ of the particular group (Faase & Pujdak 1987, p. 81).

It is easier for an internal evaluator to meet these requirements, although a skilled external evaluator will be able to develop sensitivity over time. In some cases, such as being ‘attuned’ to the context of a religious group, it may be impossible for an external evaluator to be found who meets the needs or sensitivity of the group.

Ability to collect information

One potential advantage of external evaluators over internal ones is that they are more able to collect information that might be difficult to obtain (Braskamp et al. 1987).

People can be more willing to open up in the presence of a stranger. At the same time, an external evaluator is often able to help people gain a new perspective because of his or her wider perspective (Weiss 1972, p. 38). In some cases, 'politics and turf battles' will make it difficult for an internal evaluator to gain information and support (Goldberg & Sifonis 1994, p. 27).

Flexibility

A contrary claim made for internal evaluators is that they are more flexible and in a position to alter the evaluation design quickly if it becomes clear that an evaluation activity is not useful (summarised in Cummings et al. 1988, pp. 60–1). However, it is hard to see why internal evaluators would be more responsive to this feedback than external evaluators. Presuming that both are reviewing data as obtained, it should be possible for both to respond to difficulties in the evaluation process.

Specialist skills and expertise

External evaluators are most often selected because of their experience and specialist expertise (Goldberg & Sifonis 1994, p. 27; Braskamp et al. 1987). Sometimes, this can lead to organisations only being impressed by the credentials of academic researchers, ignoring valuable internal expertise (Weiss 1972, p. 37).

At the same time, evaluation expertise in the specific area being evaluated is often the most sought after quality by clients (Braskamp et al. 1987, p. 66). External evaluators may be viewed as 'too remote from the realities, too ivory tower and abstract' (Weiss 1972, p. 37).

This suggests that the ideal candidate could either be an internal evaluator with specialist skills or an external evaluator with experience in the particular setting. It is impossible to generalise in regard to this factor.

Objectivity

Many people believe that external evaluators come to an evaluation unbiased and with an open mind in contrast with internal evaluators who are part of an organisation with its own history and modes of behaviour. Cummings et al. (1988, pp. 60–1) summarise arguments that internal evaluators may find their objectivity compromised by the policies of the organisation and its underlying value system. However, this is not supported by most current evaluation theory.

For decades there have been theorists showing the flaws in an objectivist approach (House 1980, p. 252). No matter how 'neutral' or 'impartial' an evaluator attempts to be, he or she will harbour implicit values—many of which are unconscious. No evaluator can be completely 'disinterested and aloof' from the society of which he or she forms

part (p. 254). The best that can be achieved is evaluator impartiality. Furthermore, a number of radical critiques doubt the ability of any evaluator to provide an 'objective stance' (Kushner 2000) and record reality (Guba & Lincoln 1987). Fewer authors, such as Scriven (1997), argue for objectivity as a continuing ideal.

Noting that 'evaluation has not yet repositioned itself in relation to influential debates stimulated by post-modernism and post-structuralism' (p. 10), Kushner argues that evaluation methodology necessarily privileges one view of the world over another—usually selecting the view of program administrators, funding bodies and other powerful actors over that of the individuals who participate in programs. Because evaluation methodology pursues a logic of 'coherence' (p. 13), it necessarily selects a point of view that will deny meaning to other 'different worlds of meaning' involved.

Guba & Lincoln (1987) trace the 'four generations' of evaluation through the roles of technical measurement, description, judgement to the current role of negotiator or change agent (pp. 203–9). According to this analysis, evaluation involves deciding between the claims, concerns and issues put forth by members of a variety of

Many people believe that external evaluators come to an evaluation unbiased and with an open mind in contrast with internal evaluators who are part of an organisation with its own history and modes of behaviour.

stakeholder groups—exercising value pluralism in evaluation work. By asserting the worth of one outcome over another, evaluators become involved in a struggle of contexts and values. Evaluation does not record, but rather creates 'reality' (p. 217).

Because of the subjectivity of the values involved, there is no reason to suppose that internal or external evaluators would be any more or less likely to share the values or 'reality' of the people involved in the evaluation.

The individuals employed by an organisation are not substantially more likely to share 'value systems' with each other than outsiders. Value systems are complex and made up of a myriad of factors including culture, upbringing, experience and education.

'Objectivity' should not be a factor in selecting internal or external evaluators.

Perceived objectivity

By contrast, perceived objectivity can be an important factor in choosing between an internal and external evaluator.

One of the things that many people look for from an evaluation is credibility or perceived objectivity (Braskamp et al. 1987; Love 1991, p. 5; Mathison 1994, p. 301). Free from the post-modern challenge to objectivity, many users of evaluation are looking for an *appearance* of objectivity—and will be satisfied

with an independent evaluation by someone with 'no obvious stake in the program' (Weiss 1972, p. 37).

Unless there is some prior relationship between the organisation and the external evaluator, the external evaluator will be presumed to be unbiased and 'objective.'

By contrast, an internal evaluator will need to rely on standards such as 'professional competence, objectivity, and clarity of presentation' (Braskamp et al. 1987, p. 65). In most cases, these will be sufficient and a transparent methodology will allow the results to speak for themselves.

While both external and internal evaluators have their biases (Love 1991, p. 5), the appearance of impartiality remains a real need among many evaluation audiences and remains a strong argument for the use of external evaluators, especially in contested areas where internal evaluators might be seen as politically motivated or biased. Even so, there remains a risk that internal evaluators will not be seen as objective and reliable as external evaluators. Weiss (1972, p. 38) calls this a 'lingering suspicion'.

Organisations should take the issue of perceived impartiality very seriously in selecting evaluators. In some cases this will be a strong factor for nominating an external evaluator.

Accountability for use of government funds

The case for perceived objectivity of evaluation becomes even stronger for a government program.

According to House (1980), evaluation which is undertaken for private purposes needs to be strongly distinguished from a public evaluation. Instead of merely affecting the private recipient, the evaluation is now fundamentally and inextricably part of a *public* situation (House 1980, p. 18). This places certain constraints on the evaluation and suggests the desirability of a transparent evaluation process.

Public demand for government accountability was one of the main reasons for the growth of evaluation studies (Meyers 1981, p. 2) and remains a key part of the idea of reinventing government (Patton 1997, p. 14). Evaluation is one of the key ways of legitimising government in modern capitalist societies (House 1993, p. 32).

While transparent evaluation can be achieved by both internal and external evaluators, the legitimisation achieved is likely to be greater with an external evaluator. This should be considered as a factor by government in selecting evaluators and by organisations receiving government funding.

Willingness to criticise

Another argument for external evaluators is that, as new unbiased actors, they may be more forthright about their recommendations, refusing to be 'buffaloed' and daring to 'scare the hell out of people' (Braskamp et al. 1987, p. 65). Weiss (1972, p. 38) notes that external evaluators can often raise issues that would be uncomfortable for an internal evaluator to raise.

There can be negative consequences for an internal evaluator, both professional and social, if he or she criticises colleagues (Weiss 1972, p. 38). Specifically, a negative evaluation may have implications for the careers of program staff and jeopardise the organisation's chances of future funding (Owen with Rogers 1999, p. 140). There may be pressure on an internal evaluator to become a 'public relations tool' of the organisation (Cummings et al. 1988, p. 60–1). According to Scriven 1997 the role of internal evaluator is difficult because the program is not only 'paymaster' but is also the 'social environment in which the evaluator has worked for some time' (p. 487).

At the same time, external evaluators are often under the same pressures to give favourable assessments as internal evaluators (Weiss 1972, p. 38). Many of the same issues of manipulation and threat arise, such as where an external evaluator has hopes of additional work or whose payment might be withheld (Mathison 1994, p. 301). Generally, external consultants have less security and legal protection than staff.

In addition, if an external evaluator takes a highly participative approach to the evaluation, it may become difficult to criticise the people who the evaluator has taken such pains to cultivate (Scriven 1997; Patton 1997, p. 44). Social pressures on external evaluators can be as intense as those on internal staff (Weiss 1972, p. 38). This should not be a major factor in most organisations' decision about which type of evaluator to employ.

Utilisation of evaluation

Utilisation of evaluation is a key issue for many practitioners (Posavac & Carey 1997; Patton 1997). Some argue that an internal evaluator is better placed to understand the environment and prepare findings in the style most likely to be used (Shapiro & Blackwell 1987). An internal evaluator can also build credibility over time (Cummings et al. 1988, pp. 60–1) and prepare the ground for acceptance and utilisation of evaluation results (Gunn 1987, pp. 15–16, Love 1991, p. 5). Because internal evaluators know the 'nuances' of the organisation, they are able to see ways that the evaluation can make a difference and promote the use of evaluation findings over the longer term (Weiss 1972, pp. 37–9).

The ability to communicate relevant and timely information is a form of credibility that may be as much of an advantage for an internal evaluator as the 'objectivity' of external evaluators (Love 1991, p. 4). However, many authors believe that this issue can be solved by external evaluators working closely with stakeholders in a participative mode (Posavac & Carey 1997).

The issue of utilisation becomes particularly relevant for utilisation of information for organisational change. If one of the aims of evaluation is to 'focus on creating a common language for improvement efforts' (Czarnecki 1999), it makes sense to have internal evaluators who can

use ongoing measurement processes to continue to generate improvements. Internal evaluators can be particularly useful in 'institutionalizing the monitoring process' and building systems that are able to respond quickly to address performance issues (Goldberg & Sifonis 1994, p. 237–8).

These factors are most relevant for evaluations which are to be used for internal improvement processes, although they do have some relevance for external dissemination. They may play a small, or determining, part in choosing an evaluator depending on the particular facts of the case.

Dissemination of results

Dissemination of results is also an important issue for evaluators (Rossi & Freeman 1982). Cronbach et al. 1981 point out that a successful evaluation must do more than just 'amass good data': timely communications should 'distribute information to the persons rightfully concerned, and those hearers should take the information into their thinking' (pp. 65–6). All evaluators should be aware of communication issues.

This factor should not favour internal evaluators or external ones.

Ethical issues

Both internal and external evaluators face a number of ethical issues. Internal evaluators arguably have to deal with more stark cases of divided loyalty and pressure to suppress negative results (Love 1991, p. 59); however, external evaluators face similar issues.

There is no compelling ethical reason to prefer external over internal evaluators.

Organisational investment

Finally, the availability of an internal evaluator could be seen as an investment that an organisation makes in 'an enduring corporate resource' that is useful in a number of contexts (Love 1991, p. 5).

Whether this is a reasonable investment will depend on the size of the organisation and its likely future evaluation needs. The argument for building staff capacity may be a strong factor in some cases but may make no sense for an organisation that is only going to conduct one evaluation per decade. This factor will depend on the particular context of each case.

Guidelines for decision-making

Many of the factors outlined above are finely balanced. The following checklist (Table 1) gives some guidance for organisations deciding between internal or external evaluators.

'Weak' indicates that the factor slightly favours one type of evaluator over the other while 'strong' indicates that it is more likely to be a determining factor (see table below).

While this table offers some guidance, it is important to be aware that both internal and external evaluators are able to fulfil a wide range of roles. Whether an evaluator adopts the evaluation roles such as 'consultant', 'facilitator' or 'director' of evaluation may have more influence on the factors above than whether the evaluator is internal or external (see Weiss 1972, p. 99, Newman and Brown 1996, p. 123, and Rothwell and Cookson 1997, p. 14 for a discussion of roles and role acquisition).

TABLE 1: CHECKLIST FOR DECIDING BETWEEN INTERNAL OR EXTERNAL EVALUATORS

FACTOR	INTERNAL	EXTERNAL	GUIDELINE
Cost	Weak		Cost comparison calculation needed in each case
Availability	Weak		Assessment of current availability needed in each case
Knowledge of program and operations	Weak		Depends on amount of organisational information
Knowledge of context	Weak		Depends on how unusual the organisation is
Ability to collect information		Weak	Depends on how 'territorial' the organisation is
Flexibility			Not a determining factor
Specialist skills and expertise			Not a determining factor
Objectivity			Not a determining factor
Perceived objectivity		Strong	May be important for 'sensitive' evaluations and specific audiences
Accountability for use of government funds		Strong	Government and organisations receiving government funding should consider this factor
Willingness to criticise		Weak	Not usually a determining factor
Utilisation of evaluation	Weak		Depends on purpose of evaluation, especially if focused on organisational improvement
Dissemination of results			Not a determining factor
Ethical issues			Not a determining factor
Organisational investment	Weak		Depends on organisation's future evaluation needs

When drawing conclusions about the qualities of internal and external evaluators, it is important to note the variation caused by these evaluation roles. For example, an external evaluator working in 'facilitator' mode may be better able to prepare for utilisation of results than an internal evaluator working in 'director' mode. This analysis is implied in Scriven (1997, pp. 487–9) who suggests that bias can occur where an external evaluator is hired to assist with advocacy or training. The role that the evaluator chooses to play in the evaluation may be more important in that situation than whether the evaluator is internal or external.

It should also be noted that there is a third option available: making use of both internal and external evaluators in 'partnership' to gain the benefits of both approaches. This model is a hybrid offered by Shapiro and Blackwell (1987) and Mathison (1994, p. 300) among others and involves sharing the evaluator's role with an external consultant, dividing tasks such as collection of data and data analysis.

Such an arrangement lowers costs while conferring some of the credibility benefits of external evaluation. Another advantage would be the role of the external evaluator as an 'evaluation educator' (p. 60), building an organisation's skills and knowledge as an investment for the future. This may be an additional option in some cases.

Conclusion

Despite the tone of much of the evaluation literature surveyed in the first part of this paper, evaluation by internal evaluators is not a radical phenomenon. One estimate places internal evaluations as three quarters of all evaluation activities in North America (Love 1991, p. 2).

The debate about choosing between external and internal evaluators should be seen as part of the broader question about whether evaluation should be considered as a profession or as a modality. While there is no scope in this paper to discuss this issue fully, a few concluding thoughts are offered here.

There are two different world views apparent in the literature reviewed: in one world view, evaluation is something that should be carried out primarily by professionals (external evaluators), while in the other world view, evaluation skills should be spread as widely as possible (such as to management and other staff).

This relates to a long-standing debate about evaluation as a professional discipline.

Many writers have felt the need to raise the status of evaluation by putting boundaries around the profession. For example, Cronbach et al. (1981) stress the need for 'a strong professional community of evaluators who share a common vision of their work' as responsible to the 'larger social interest' (pp. 71–2). Other authors such as Rossi et al. (1999), House (1993) and Meyers (1981) also discuss a professional evaluation community. However, in House's view, evaluation still falls short of the three claims of a profession: scientific base, validation

by a community of peers and orientation towards important social values (House 1993, p. 19).

Much of the discussion of evaluation as a profession revolves around appropriate education and training: for example, Cronbach et al. (1981) argue for improvement in the education of evaluators to build interest in evaluation as an area of inquiry, not just 'an ignoble form' of other research (p. 340).

It is possible that the need to cordon off space is true for any emerging profession. Relatively recent examples such as the emergence of professional groups such as psychologists and mediators shows how this process can take place.

These examples also show, however, that the most effective strategy in the long-term is to have the emerging discipline recognised and included in mainstream institutions. While in the short-term, this may lead to reduced need for 'professionals' in the emerging discipline, in the longer term there will be increased demand for specialist skills if they are an accepted part of society.

This view of the development of evaluation would argue for greater encouragement of internal evaluators as an important part of the growth of acceptance of the legitimacy of evaluation. Using this approach, perhaps the best way to view evaluation would be not as a profession but as a 'transdiscipline' like logic or statistics: 'a discipline in its own right, but one that services other disciplines and is an integral part of other disciplinary endeavours' (House 1993, p. 86).

Ultimately, the 'Great House' of evaluation should have room for both internal and external evaluators if it wishes to continue to grow.

Note

- 1 The terminology of 'internal and external' evaluators is intended to be consistent with the definition quoted in Owen with Rogers (1999) of an insider as 'an entity (individual or group) directly associated with the conduct and/or impact of a program' and an outsider as 'an entity external to the program' [Owen with Rogers (1999, p. 135) quoting Hogben 1977]. The question has been formulated as whether the evaluator should be internal or external instead of, for example, whether the evaluator should be 'independent' to avoid confusion. 'Independent' is not a useful term in this context since both internal and external evaluators can be 'independent' depending on the evaluation role they choose. Issues of objectivity and independence are discussed in part two of this review.

References

- Braskamp, LA, Brandenburg, DC & Ory, JC. 1987, 'Lessons about clients' expectations, in J Nowakowski (ed.), *The client perspective on evaluation: New directions for program evaluation*, no. 36, Jossey-Bass, San Francisco.
- Cronbach, LJ, Ambron, SR, Dornbusch, SM, Hess, RD, Hornik, RC, Phillips, DC, Walker, DF & Weiner, SS 1981, *Towards reform of program evaluation: aims, methods, and institutional arrangements*, Jossey-Bass, San Francisco.

- Cummings, OW, Nowakowski, JR, Schwandt, TA, Eichelberger, RT, Kleist, KC, Larson, CL & Knott, TG 1988, 'Business perspectives on internal/external evaluation', in JA McLaughlin, LJ Weber, RW Covert & RB Ingle (eds), *Evaluation Utilization*, vol. 39, pp. 59–74.
- Czarnecki, MT 1999, *Managing by measuring: how to improve your organization's performance through effective benchmarking*, AMACOM, New York.
- Faase, TP & Pujdak, S 1987, 'Shared understanding of organizational culture', in J Nowakowski (ed.), *The client perspective on evaluation: New directions for program evaluation*, no. 36, Jossey-Bass, San Francisco.
- Globerson, A, Globerson, S & Frampton, J 1991, *You can't manage what you don't measure: control and evaluation in organizations*. Gower Publications, Aldershot.
- Goldberg, B & Sifonis, JG 1994, *Dynamic planning: the art of managing beyond tomorrow*, Oxford University Press, New York.
- Guba, EG & Lincoln, YS 1987, 'The countenances of fourth-generation evaluation', in DJ Palumbo (ed.), *The politics of program evaluation*, Sage Publications, Newbury Park, California.
- Gunn, WJ 1987, 'Client concerns and strategies in evaluation studies', in J Nowakowski (ed.), *The client perspective on evaluation: New directions for program evaluation*, no. 36, Jossey-Bass, San Francisco.
- Hogben, D 1977, 'Curriculum evaluation: by whom, for whom?', paper presented at the Annual Conference of the American Educational Research Association.
- House, ER 1980, *Evaluating with validity*, Sage Publications, Beverly Hills, California.
- House, ER 1993, *Professional evaluation: social impact and political consequences*, Sage Publications, Newbury Park, California.
- Kushner, S 2000, *Personalizing evaluation*, Sage Publications, Thousand Oaks, California.
- Love, AJ 1991, *Internal evaluation: building organizations from within*, Sage Publications, Newbury Park, California.
- Mathison, S 1994, 'Rethinking the evaluator role: partnerships between organizations and evaluators', *Evaluation and Program Planning*, vol. 17, no. 3, pp. 299–304.
- Meyers, WR 1981, *The evaluation enterprise: a realistic appraisal of evaluation careers, methods, and applications*, Jossey-Bass, San Francisco.
- Newman, DL & Brown, RD 1996, *Applied ethics for program evaluation*, Sage Publications, San Francisco.
- Nowakowski, J 1987, 'Editor's notes', in J Nowakowski (ed.), *The client perspective on evaluation: New directions for program evaluation*, no. 36, Jossey-Bass, San Francisco.
- Owen, JM with Rogers, PJ 1999, *Program evaluation: forms and approaches*, 2nd edn, Allen & Unwin, St Leonards, NSW.
- Patton, MQ 1997, *Utilization-focused evaluation: the new century text*, 3rd edn, Sage Publications, Thousand Oaks, California.
- Perloff, R 1979, 'Evaluator intervention: the case for and against', in R Perloff (ed.), *Evaluator interventions: pros and cons*, Sage Publications, Beverly Hills, California.
- Posavac, EJ & Carey, RG 1997, *Program evaluation: methods and case studies*, 5th edn, Prentice Hall, Upper Saddle River, New Jersey.
- Rossi, PH & Freeman, HE 1982, *Evaluation: a systematic approach*, 2nd edn, Sage Publications, Thousand Oaks, California.
- Rossi, PH, Freeman, HE & Lipsey, MW 1999, *Evaluation: a systematic approach*, 6th edn, Sage Publications, Thousand Oaks, California.
- Rothwell, WJ & Cookson, PS 1997, *Beyond instruction: comprehensive program planning for business and education*, Jossey-Bass, San Francisco.
- Scriven, M 1997, 'Truth and objectivity in evaluation', in E Chelimsky & WR Shadish (eds), *Evaluation for the 21st century: a handbook*, Sage Publications, Thousand Oaks, California.
- Shapiro, JZ & Blackwell, DL 1987, 'Large-scale evaluation on a limited budget: the partnership experience', in J Nowakowski (ed.), *The client perspective on evaluation: New directions for program evaluation*, no. 36, Jossey-Bass, San Francisco.
- Sluyter, GV 1998, *Improving organizational performance: a practical guidebook for the human services field*, Sage Publications, Thousand Oaks, California.
- Weiss, CH 1972, *Evaluation research: methods for assessing program effectiveness*, Prentice Hall, Englewood Cliffs, New Jersey.
- Weiss, CH 1998, *Evaluation: methods for studying programs and policies*, 2nd edn, Prentice Hall, Upper Saddle River, New Jersey.