Evaluation Journal of Australasia, Vol. 10, No. 1, 2010, pp. 8–16

Developing evaluation capacity and use in the New Zealand philanthropic sector:
What can be learnt from the US experience?

Michael Blewden

Michael Blewden is a Senior Researcher at Social Health Outcomes Research and Evaluation (SHORE), SHORE and Whariki Research Centre, School of Public Health, Massey University, Auckland, and is completing a PhD in Evaluation at Massey University.
Email: <M.B.Blewden@massey.ac.nz>

Growth in the use of evaluation within the United States (US) philanthropic sector since the late 1990s has been driven by an increasing focus on program outcomes and increased demands for evidenced-informed decision-making. With similar developments now apparent within the New Zealand philanthropic sector, it is timely to consider what has been learnt from the US experience and how this may inform local efforts to build evaluation use and capacity. Recent literature on the US experience is reviewed to identify drivers of US developments, progress made, issues arising, and lessons learnt. With attention to the current stage of capacity building within the New Zealand sector, implications and direction for local development are considered.

Introduction

While formal evaluation is not always feasible, or necessary to guide philanthropic decision-making (Mark & Beery 2004), grantmakers in the United States (US) have always used some form of evaluation to guide their work (Mark & Beery 2004). Evaluation has been used to inform funding allocation (Patrizi & McMullan 1998), for contract management (Hoff 1999; Mark & Beery 2004; McNelis & Bickel 1996), for output monitoring (Porter & Kramer 1999), and to demonstrate theories of change (Behrens & Kelly 2008).

While evaluation has historically been used less readily to assess program performance and outcomes (Kramer & Bickel 2004; Mark & Beery 2004; McNelis & Bickel 1996; Patrizi & McMullan 1998; Porter & Kramer 1999; Slater, Constantine & Braverman 2004; Wisely 2002), evidence suggests this situation is changing (Carman & Fredericks 2008; Ostrower 2004), prompted it would seem by increasing demand for evidence-informed decision-making and an outcomes focus (Buckmaster 1999; Chelimsky 2001; Hamilton et al. 2005; Patrizi & McMullan 1998; Porter & Kramer 1999; Slater, Constantine & Braverman 2004).

The recent establishment of the Evaluation Collaborative Project (ECP) in New Zealand reflects similar evaluation developments within the New Zealand philanthropic sector. It seems timely therefore to review the US experience in building evaluation capacity and use, and how this may guide New Zealand developments.

Recent literature is reviewed to summarise US developments and to reflect on progress made, issues and challenges arising. Suggested directions for local development are considered.

The term ‘grantmaker’ is used in this article to describe collectively the range of trusts, foundations, businesses and personal donations, bequests and legacies that comprise sources of philanthropic funding. Meanwhile, groups that receive grants from grantmakers are collectively referred to as 'recipients' and include non-profit clubs, associations and organisations, as well as public charities.

The US philanthropic sector

The size of the US philanthropic sector is immense. In 2001, there were 61,810 grantmaking foundations in the US, controlling an estimated $477 billion in assets and distributing grants approximately valued at $30.5 billion (Foundation Center 2003). Millet (2006) estimates that there are approximately 70,000 US foundations, although notes that only 4000 of these are staffed by two or more people. In 2004, public charities reported total revenues of nearly $1.1. trillion, and the non-profit sector as a whole accounted for 8.3 per cent of all wages and salaries paid in the US (Urban Institute 2010).

Despite its importance, the US philanthropic sector has historically made limited use of evaluation to assess program performance and outcomes (Carter 1992; Hoff 1999; McNelis & Bickel 1996; Patrizi & McMullan 1998; Porter & Kramer 1999; Slater, Constantine & Braverman 2004; Wisely 2002). Indeed in the early '90s, evaluation use in the sector was described as 'fragmentary' (Carter 1992, p. 33), and there existed significant gaps in grantmakers' evaluation knowledge (McNelis & Bickel 1996). Porter and Kramer (1999) went on to report that grantmakers devoted 'little effort … to measuring results' (p. 122), often saw performance measurement as 'unrelated to their charitable mission' (p. 122), and rarely ascertained the degree of program success.

High levels of autonomy and limited external accountability are considered two reasons why US grantmakers haven't made greater use of evaluation in the past (Hall 2004; Kramer & Bickel 2004; Leviton & Bass 2004). Other barriers to use have included a limited understanding of evaluation methods and approaches (Carter 1992), limited numbers of philanthropic staff focused on evaluation (Kramer & Bickel 2004), perceived high cost (Carter 1992; Gray 1993), and stakeholder dissatisfaction with the relevance, timeliness and usefulness of evaluation (Patrizi & McMullan 1998).

Drivers of increased attention to evaluation

Despite such history, greater attention to evaluation can be observed within the US philanthropic sector since the late 1990s. This is commonly attributed to three main factors: increasing pressure from stakeholders for more effective and accountable philanthropy (Behrens & Kelly 2008; Buckmaster 1999; Chelimsky 2001; Hamilton et al. 2005; Patrizi & McMullan 1998; Porter & Kramer 1999; Slater, Constantine & Braverman 2004); government legislation designed to hold non-profit organisations more accountable for outcomes (Carman, Fredericks & Introcaso 2008); and an initiative undertaken by United Way America (UWA), the largest philanthropic organisation in the US, to measure the outcomes of its work (Behrens & Kelly 2008; Hendricks, Plantz & Pritchard 2008). Each driver is briefly discussed.

More effective and accountable philanthropy

Concern about a seeming lack of accountability and the limited use of evaluation within the US philanthropic sector grew on a number of fronts during the 1990s. Scriven (1994) was an early critic, challenging what he termed the 'do-gooder fallacy' (p. 6), that is, the belief that in undertaking worthwhile activities with good intentions, philanthropic evaluation was unnecessary. Porter and Kramer (1999) also critiqued the sector for being satisfied with an 'historic agenda of doing good' (p. 122) while having little understanding of program effectiveness and outcomes. They called for more strategic grantmaking, the development of long-term relationships with grantees that built their capacity, and greater attention to understanding and improving the effectiveness of grantmaking.

Calls for greater accountability also grew out of recognition that compared with other organisations, grantmakers had relative independence and flexibility to be innovative in addressing social problems (Behrens & Kelly 2008; Bickel, Millett & Nelson 2002; Carter 1992; Chelimsky 2001; Kramer & Bickel 2004; Porter & Kramer 1999; Slater, Constantine & Braverman 2004). Grantmakers were thus seen as having a particular obligation to account for decision-making and funding outcomes, to use evaluation to enhance program and organisational effectiveness, and to share learnt knowledge (Bickel, Millett & Nelson 2002; Kramer & Bickel 2004).

Legislation and government mandate

After a period of reduced US federal expenditure on social services and evaluation during the late '70s and early '80s, the use of evaluation and performance management of federally-funded non-profit organisations increased in the late '80s (Carman, Fredericks & Introcaso 2008). While typically taking the form of input and output monitoring, this tradition of performance measurement is regarded as having provided the foundation for the later growth of philanthropic evaluation (Carman, Fredericks & Introcaso 2008). The early 1990s also saw a strong federal focus on results and accountability across a range of sectors, reflected in tools such as the *Government Performance and Results Act* (GPRA), which was introduced in 1993 (Behrens & Kelly 2008; Carman, Fredericks & Introcaso 2008), and the more recent Program Assessment Rating Tool (PART) (Office of Management and Budget 2004).

Outcome measurement by United Way

The UWA is a national organisation supporting more than 1300 local United Way (UW) agencies across the US (Hendricks, Plantz & Pritchard 2008). In 1995, the UWA implemented a methodology for evidencing outcomes from UWA programs (Behrens & Kelly 2008). In measuring and tracking outcomes, the initiative represented a major shift from input and output monitoring while also avoiding more expensive and complex evaluation approaches (Hendricks, Plantz & Pritchard 2008).

Progress made

The literature provides a contradictory picture of progress made in developing evaluation use and capacity within the US philanthropic sector. Greater use of evaluation is reflected in increasing numbers of full-time evaluation directors within foundations as well as new evaluation units (Patrizi & McMullan 1998). In fact, a 2004 study reported a majority of large, and a significant proportion of small, foundations using formal evaluation (Ostrower 2004). One estimate put US foundations spending over $100 million each year on evaluation (Kramer & Bickel 2004). Numerous resources now exist in support of evaluation use (see for example, Braverman, Constantine & Slater 2004; Chinman, Imm & Wandersman 2004; Plantz, Greenway & Hendricks 1997; Urban Institute 2004; Urban Institute & The Centre for What Works 2006; WK Kellogg Foundation 1998), as do a number of organised collations and networks.[1]

However, observers have continued to question the appropriateness, effectiveness and utility of philanthropic evaluation (Bickel, Millett & Nelson 2002; Guthrie et al. 2005; Hoefer 2000; Kramer & Bickel 2004; Patrizi 2006; Patrizi & McMullan 1998; Wisely 2002). Behrens and Kelly (2008) report an ongoing struggle to match growing expectations for accountability and evaluation with the sectors' capacity to design, conduct and use evaluation effectively. Dissatisfaction with the usefulness and relevance of evaluation among grantmakers continues (Braverman, Constantine & Slater 2004; Millett 2006; Patrizi 2006).

Issues and challenges

What then are the issues and challenges emerging from the US experience? Five main themes identifiable in the literature are discussed.

Relationship/understanding of evaluation

Negative preconceptions and misunderstanding about evaluation is considered to have led to tensions within the sector about how, when and why evaluation should be used (Leviton & Bass 2004; Patrizi & McMullan 1998). The use of inappropriate methods and unrealistic expectations about what evaluation can achieve is also seen to have contributed to disappointment and frustration with evaluation (Southern, cited in Behrens & Kelly 2008; Kramer et al. 2007).

Evaluation approach

Some observers have attributed stakeholder disappointment with evaluation to the use of 'traditional outcome evaluation'[2], because it is seen as having positioned evaluation as an accountability mechanism, focused primarily on reviewing outcomes from past grants retrospectively and having attempted to 'prove' attributable impacts from specific grants (Kramer et al. 2007; Millett 2006). Kramer et al. (2007) also identify numerous difficulties with this approach within the philanthropic context. They assert that scientific methods 'rarely prove that a foundation initiative "worked" to solve a problem in a replicable way' (p. 13). The premise that a lasting 'solution' can be discovered is considered particularly inappropriate when addressing complex social problems. Similarly, Millett (2006) believes that much of the work undertaken by grantmakers cannot be placed within the conventional 'attribution' framework easily (p. 8). Millet feels the expectation that evaluation will provide 'silver bullet' (p. 8) answers to questions about attribution and impact as having diverted attention from how evaluation can be used most effectively to inform and provide learning throughout program development and implementation.

Meanwhile, Kramer et al. (2007) report sector dissatisfaction with traditional outcome evaluation and so encourage grantmakers to explore simpler, alternative and less costly evaluation approaches, and those more appropriate to social innovation. They describe a 'profound transformation' (p. 9) taking place, with evaluation increasingly used as a 'forward-looking enterprise' (p. 9). Evaluation in this case is less concerned with attribution and more concerned with utility through providing 'current information and actionable insights' (p. 9), and understanding of what is needed to increase program effectiveness.

Limited learning

While there is evidence of increased levels of evaluation activity within the US philanthropic sector (Ostrower 2004), low resultant utility in terms of organisational learning and program improvement has been widely reported (Behrens & Kelly 2008; Campbell & McClintock 2002; Guthrie et al. 2005; Hendricks, Plantz & Pritchard 2008; Hoole & Patterson 2008; Millett 2006; Snibbe 2006; York 2003). Observers have noted excessive data collection and marginal data utility (Snibbe 2006), poor-quality evaluation (Behrens & Kelly 2008), and evaluation positioned as an external, expert driven, and measurement-focused 'tick the box' exercise, tolerated by organisations to secure funding (Kramer et al. 2007; Millett 2006; Patrizi 2006). Factors identified as contributing to this situation include grantmakers and recipient groups having to deal with competing evaluation frameworks and data collection requirements (Behrens & Kelly 2008; WK Kellogg Foundation 1998), accountability pressure on funders, and funders who demand accountability while doing little to build the learning capacity of recipients (Hoole & Patterson 2008).

In response to such concerns, and in recognition that grantmakers and grantees are obliged to learn from their activities (Porter & Kramer 1999), there is increasing use of philanthropic evaluation as a tool for knowledge management and learning (Behrens & Kelly 2008; Campbell & McClintock 2002; Chin 2006; Grantmakers for Effective Organisations 2005, 2007; Guthrie et al. 2005; Hernandez & Visher 2001; Kramer et al. 2007; Millett 2006; Patrizi & McMullan 1998; Putnam, 2004; Woodwell 2005; York 2003). Learning-oriented evaluation is focused on inquiry (Patrizi 2006) and is used to 'improve', not just 'prove' (Hoole & Patterson 2008; York 2003); it seeks to understand how and why change occurs (Behrens & Kelly 2008; Conner et al. 2004), and is focused on strengthening programs and services and sharing learning (Conner et al. 2004; York 2003). A learning approach reframes evaluation as a process of evaluative inquiry that integrates evaluative thinking through all levels and stages of grantmaking (Patrizi 2006).

Millett (2006) advocates 'contribution and learning' (p. 8), rather than attribution, as a more 'reasonable and practical standard' (p. 8) for philanthropic evaluation. He calls for a more equitable and functional partnership between grantmakers and grantees that recognises grantee effectiveness as being critical to success and which supports evaluation as a means to learn and optimise program effectiveness.

The 'prospective approach to policy change evaluation' (Guthrie et al. 2005, p. 5) put forward by The California Endowment can be seen as an example of what Millet is looking for. The approach recognises that attributing specific contributions from any individual player or intervention is often impossible and meaningless within a change environment where success requires multiple change agents, strategies and activities. Evaluation is positioned as a collaborative activity that begins at project onset, is responsive to mid-course adjustments, and which provides ongoing monitoring and assessment. Monitoring and impact indicators, developed from theory-of-change models, are used to provide information on meaningful contributions, and progress towards change goals, rather than evidence of individual attribution. The approach emphasises learning and building grantees' self-evaluation capacity.

Increasing diversity in evaluation approaches

Philanthropic evaluation is increasingly drawing upon a range of approaches (Hall 2004; Kramer & Bickel 2004; Kramer et al. 2007; Slater, Constantine & Braverman 2004) and is reflective of practitioners' diverse training and values (Kramer & Bickel 2004). It is recognised that no single approach to evaluation will work best in all situations (Kramer et al. 2007; Patrizi & McMullan 1998; Putnam 2004), and diverse approaches are thus appropriate. However, in association with this, grantmakers are faced increasingly with having to match needs and context with different approaches (Behrens & Kelly 2008; Kramer et al. 2007; Slater, Constantine & Braverman 2004), but this requires understanding of the possibilities and limitations of different evaluation options in different circumstances (Evaluation Roundtable, cited in Behrens & Kelly 2008; Hall 2004; Kramer & Bickel 2004; Kramer et al. 2007). Without such knowledge, it has been asserted that grantmakers may struggle to be discerning clients of evaluation, with appropriate expectations and the ability to match skills, approaches and needs (Kramer & Bickel 2004).

There are calls, therefore, for initiatives that communicate different evaluation methods and approaches and which build the capacity of grantmakers to select and use evaluation appropriately (Evaluation Roundtable, cited in Behrens & Kelly 2008; Hall 2004; Kramer & Bickel 2004; Kramer et al. 2007). Also identified is the need for a systematic approach to determining what to evaluate and at what intensity, taking into account factors such as the size of grantmaker or grant, and the decision-making need (Mark & Beery 2004).

Systemic, organisational and cultural barriers

Systemic, organisational and cultural barriers to evaluation within the US philanthropic sector are also discussed in the literature (Bickel, Millett & Nelson 2002; Kramer & Bickel 2004; Millett 2006). Grantmakers can be highly autonomous and can face limited external accountability in regard to outcomes (Hall 2004; Kramer & Bickel 2004; Millett 2006). Consequently, it has been asserted that grantmakers often work within uncritical environments, can presume recipients to be well intentioned, and can be reluctant to enquire too critically into performance and outcomes (Bickel, Millett & Nelson 2002). Within such a context, honesty and negative evaluation findings may be particularly unwelcome, and learning from evaluation difficult (Kramer & Bickel 2004). Further, grantmakers may develop inflated expectations about outcomes, with recipient groups subsequently 'over promising' in order to secure funding (Bickel, Millett & Nelson 2002; Kramer & Bickel 2004; Millet 2006). Grantmakers can see themselves as social experimenters and already 'learning organisations' (Bickel, Millett & Nelson 2002; Millett 2006), thus supporting a willingness to take, but not learn from, risks and/or resultant failure.

Furthermore, it has been noted that grantmakers are often required to process a high volume of grants while also reducing costs, and that these combined pressures can undermine evaluation use (Kramer & Bickel 2004; Millett 2006). Evaluation can look particularly expensive when funds are scarce and less necessary when funds are plentiful (Bickel, Millett & Nelson 2002; Millett 2006). A cost-cutting and 'just do it' culture may also undermine capacity, resources or incentive to build evaluation capacity (Bickel, Millett & Nelson 2002).

It has also been observed that there is often a high level of professional autonomy among grantmaking staff (Kramer & Bickel 2004; Millett 2006), and therefore some reluctance for there to be a right or wrong way of doing things. Evaluation may be discouraged if programs evaluated are disadvantaged through being compared by decision-makers with programs relying on more superficial, and probably, more positive assessment (Bickel, Millett & Nelson 2002). Staff-level discouragement is more likely if decision-makers are not driving the evaluation and learning effort (Bickel, Millett & Nelson 2002). This is suggested to be a common scenario (Kramer & Bickel 2004).

Learning from the US experience

So, what might be learnt from the US experience? And how might this direct the Evaluation Collaborative Project (ECP)—a joint initiative of Philanthropy New Zealand (PNZ), the Lottery Grants Board (LGB) and ASB Community Trust (ASBCT)—to build evaluation capacity and use it across the philanthropic, community and voluntary sectors in New Zealand (Philanthropy New Zealand 2008)?

New Zealand philanthropic sector

The total operating expenditure of paid and voluntary effort within the New Zealand non-profit sector[3] has been estimated at NZ$9.8 billion for the year ending March 2004 (Sanders et al. 2008). The sector as a whole employs over 200 000 full-time equivalent paid staff and volunteers, representing 9.6 per cent of the economically active workforce (Sanders et al. 2008). Approximately half (49 per cent) of the non-profit workforce is engaged in expressive activities[4] reflecting the significant presence of cultural, recreational and sporting associations, and Maori civic organisations. Over three-quarters of total philanthropic grant funding is distributed across culture, sport and recreation, education and research, social services, and health (Slack & Leung-Wai 2007). Service activities[5] account for 50 per cent of total activity in the non-profit sector, a small proportion compared with other countries and reflecting a high level of direct government funding to state institutions to provide universal public provision of such services (Sanders et al. 2008).

Sources of philanthropic giving in New Zealand include trusts and foundations, individuals and businesses, with trusts and foundations accounting for the greatest proportion of total giving (Slack & Leung-Wai 2007). In 2004, philanthropic giving represented 20 per cent of total revenues for the New Zealand non-profit sector, with fees (55 per cent) and government (25 per cent) providing the remainder (Sanders et al. 2008). Including $65 million from community and energy trusts and $260 million from gaming machines trusts, total giving was NZ$1.6 billion in 2004, this equating to 1.1 per cent of New Zealand's GDP (Sanders et al. 2008).

Direction for local evaluation capacity building

The ECP was initiated in 2008 and has undertaken a number of evaluation workshops with PNZ members. However, the capacity building effort can be considered in its early stages, thus presenting a unique opportunity to learn from the US experience. Learning and implications are considered in respect to the overall capacity-building strategy, the development of grantmakers and recipient groups, and potential structural barriers to evaluation.

Capacity-building strategy development

The US experience has revealed that negative preconceptions and misunderstanding about evaluation can be a barrier to use, and reinforces the importance of initially building a shared vision of how evaluation can contribute to effective philanthropy from the outset. Demystifying the process and products of evaluation (Leviton & Bass 2004) and increasing stakeholders' understanding of evaluation, are indicated to be key initial tasks. Suggested focus areas include ensuring that stakeholders understand the potential and limitations of evaluation within the philanthropic context, the use of evaluation for outcome, and the use of evaluation as a learning and knowledge management tool.

The US experience indicates the importance of key decision-makers, such as board members, being involved in developing the evaluation agenda and vision (Kramer & Bickel 2004). Leadership from decision-makers helps to integrate evaluation into program and strategy decision-making (Evaluation Roundtable, cited in Behrens & Kelly 2008) and enhances evaluation utility and learning outcomes (Bickel, Millett & Nelson 2002; Kramer & Bickel 2004). The ECP has signalled a commitment to developing an 'evaluation for learning' approach and building decisions-makers' commitment to this will be important (Grantmakers for Effective Organisations 2007; Hoole & Patterson 2008; Kramer & Bickel 2004; Patrizi 2006; Patrizi & McMullan 1998; Putnam 2004; York 2003). A number of board members have attended the initial workshops held by the ECP and the ongoing involvement of such personnel may help to avoid subsequent evaluation being conducted in isolation from organisational mandate and support (Patrizi & McMullan 1998).

Given resource constraints, it is unlikely that sufficient evaluation resources will ever be available within the New Zealand sector to permit significant investment in large-scale outcome evaluation using experimental methods. Such constraints may be one factor shaping the ECP's commitment to an 'evaluation for learning' approach and the use of simple, accessible evaluation methods. While the role and place of outcome evaluation utilising experimental methods is acknowledged (Kramer et al. 2007), a greater focus on learning-oriented approaches may help avoid some of the unintentional consequences reported from the US; for example, excessive and competing data collection demands, limited data utility, poor learning and program improvement outcomes, and failure to integrate evaluation within organisations and daily activity. However, the US experience also shows that a learning-oriented approach to evaluation needs to be developed and implemented carefully. While it is beyond the scope of this article to discuss this in depth, some general principles for action can be summarised from the literature.

A learning-oriented approach requires stakeholder understanding and agreement concerning why learning is important, who should be learning, and for what purpose (Patrizi & McMullan 1998). It will likely require some compromise in the level of sophistication achievable in evaluation designs, as well as the need to balance feasibility, affordability and accountability demands (York 2003). Leaders need to give priority to evaluation for program improvement (Hoole & Patterson 2008; Putnam 2004) and engage themselves in learning, while also supporting staff to learn and apply their learning (Botcheva, White & Huffman 2002). Learning needs to be a joint responsibility, integrated into daily processes (Grantmakers for Effective Organisations 2007; Patrizi 2006; York 2003), and the building of learning capacity accepted as central to organisational capacity building (York 2003).

Grantmaker development

Similar to the USA, it seems likely that building the capacity of grantmakers to select and use evaluation appropriately will also be important locally. A range of evaluation approaches are also available to local evaluators (Davidson 2003; Duignan 2003). In addition, local practice is informed by a diverse base of training, theories, methods and values (Davidson, Lunt & McKegg 2003). Such conditions reinforce the importance of grantmakers being able to understand the possibilities and limitations of different evaluation options within different circumstances. The US experience suggests that if grantmakers understand the respective strengths and weaknesses of various approaches, they will have more realistic expectations for evaluation, and may be more successful in selecting and using evaluation appropriately.

The limited size and resources of some US philanthropic organisations have been described as barriers to evaluation (Kramer & Bickel 2004), as well as factors to consider when determining the appropriate use of evaluation (Mark & Beery 2004). Considering the relatively modest size of grants typically made in New Zealand (Philanthropy New Zealand 2008), it seems that local guidance in determining what to evaluate, and at what intensity, would also be appropriate. Grantmakers may resist evaluation if it is not clear that evaluation approaches need to account for factors such as grant size, the questions and decisions at stake, and intended use by potential users. Also a belief that evaluation costs are likely to be disproportionate to program budgets, will almost certainly feed resistance.

Recipient group development

Top-down mandates for evaluation are unlikely to support a learning approach (Bickel, Millett & Nelson 2002) and the importance of building grantees' readiness and capacity to learn is widely observed in the USA. It is recommended that grantmakers initially work with recipient groups to assess their organisational and program readiness to make evaluation a more integral part of their work (York 2003), to identify their learning goals (Grantmakers for Effective Organisations, 2007b; York, 2003), and to build their capacity to use evaluation for learning (Grantmakers for Effective Organisations 2007). Data collection and reporting requirements should be mission-driven (Grantmakers for Effective Organisations 2007; Hoff 1999; Hoole & Patterson 2008; Putnam 2004), and linked to addressing learning objectives (Patrizi & McMullan 1998), which in turn should be focused on improving performance (Hoole & Patterson 2008).

A learning approach to evaluation is likely to begin with conversations within organisations and across stakeholders to build common understanding about goals and desired outcomes and the information and learning that will support, as well as signal, the achievement of these (Grantmakers for Effective Organisations 2007; Hoff 1999; Hoole & Patterson 2008; Millett 2006; Patrizi 2006; Patrizi & McMullan 1998; Putnam 2004). Assessment of current learning capacity, existing knowledge, and knowledge sharing processes, as well as required additional resources for learning, is recommended (Grantmakers for Effective Organisations 2007; York 2003). Reframing existing information collection to align more closely with identified learning objectives may reduce the risk of evaluation being perceived and resisted as an 'add-on' task requiring additional resources (Grantmakers for Effective Organisations 2007; Guthrie et al. 2005; York 2003).

Addressing structural issues

A range of systemic, organisational and cultural characteristics of philanthropic organisations have been identified as barriers to evaluation within the US sector (Bickel, Millett & Nelson 2002; Kramer & Bickel 2004; Millett 2006). While it has been observed that barriers to utilising and learning from evaluation are typically experienced in many organisations (Chelimsky 2001; Leviton & Bass 2004; Woodwell 2005), whether, and to what extent, these barriers exist locally, and thus warrant local responses, will require further consideration.

This will perhaps be best undertaken by local stakeholders who have insider understanding of the sector.

Conclusion

Progress made and issues and challenges emerging from efforts to develop evaluation use and capacity within the US philanthropic sector have been reviewed. Lessons from the US experience have been considered and used to suggest broad directions for local efforts in order to build evaluation use and capacity within the New Zealand philanthropic sector.

The US experience indicates that attention to process and strategy is critical in building evaluation capacity and use. This article provides less understanding of the specific evaluation skills, knowledge and resources that may be appropriately developed. However, such detail may 'fall out' of initial efforts to build collective understanding, consensus and commitment to the contribution evaluation can make to effective philanthropy. This article has also provided limited analysis about whether systemic, organisational and cultural barriers to evaluation exist locally. Further examination here has been recommended, with the literature providing some direction for this. There are further gaps in the extent to which I have considered local context when interpreting the US experience. In particular, implications for building evaluation capacity and use that will deliver benefits to Maori, the Indigenous population of New Zealand, as well as other cultural minorities within the country, must be considered. These groups are key stakeholders associated with the services and programs supported by philanthropy (Sanders et al. 2008), and there is growing understanding of appropriate approaches when conducting evaluations with Maori (Cunningham 2003; Irwin 1994; Moewaka Barnes 2003; Moewaka Barnes 2009; Smith 1996, 1999).

Developments in the US were clearly driven by a mix of both 'top-down' and 'bottom-up' pressures, collectively building a mandate to use evaluation to address issues relating to accountability, performance and outcomes. Whether a similar mix of drivers exists in New Zealand also warrants further examination. To date, PNZ members appear to have played a key role in driving demand for increased evaluation knowledge and utilisation. This, as well as calls for an 'evaluation for learning' approach, suggest that initial momentum in the ECP will be driven through the project successfully assisting stakeholders to experience the positive learning outcomes achievable through the effective and appropriate use of evaluation within the philanthropic context.

Acknowledgements

Thanks go to Robyn Scott of Philanthropy New Zealand, Chloe Harwood of ASB Community Trust, and Phil Harington of the New Zealand Lottery Grants Board for their valuable review and feedback on this article.

Notes

1 Including the Evaluation Roundtable, the Centre for Effective Philanthropy, the Nonprofits and Foundations Topical Interest Group within the American Evaluation Association, the Morino Institute, and the National Committee for Responsive Philanthropy.

2 Described by Kramer et al. (2007) as the longitudinal, scientific study, which compares the beneficiary group to a randomised control group, and which assumes that the effectiveness of any intervention can only be determined 'by showing that any improvement would not have occurred, but for the grant' (p. 13).

3 Defined to include *tangata whenua* governance organisations which provide stewardship of the affairs of *iwi* (tribes).

4 Includes: culture, sports and recreation; environmental protection; civic and advocacy activities; activities of labour unions; professional associations and business leagues; and religious worship.

5 Includes: education and research; health services; social services; community development; and housing.

References

Behrens, T & Kelly, T 2008, 'Paying the piper: foundation evaluation capacity calls the tune', in J Carman & K Fredericks (eds), *Nonprofits and evaluation: New directions for evaluation, 119*, Jossey-Bass, San Francisco.

Bickel, W, Millett, R & Nelson, C 2002, 'Challenges to the role of evaluation in supporting organizational learning in foundations', 'Web extra', *Foundation News and Commentary*, vol. 43, no. 2, viewed 12 October 2007, <http://www.foundationnews.org/webextra/learningweb.htm>.

Botcheva, L, White, C & Huffman, L 2002, 'Learning culture and outcomes measurement practices in community agencies', *American Journal of Evaluation,* vol. 23, no. 4, pp. 421–434.

Braverman, M, Constantine, N & Slater, J (eds) 2004, *Foundations and evaluation: contexts and practices for effective philanthropy*, Jossey-Bass, San Francisco.

Buckmaster, N 1999, 'Associations between outcome measurement, accountability and learning for non-profit organisations', *International Journal of Public Sector Management*, vol. 12, no. 2, pp. 186–197.

Campbell, M & McClintock, C 2002, 'Shall we dance? Program evaluation meets OD in the nonprofit sector', *OD Practitioner*, vol. 34, no. 4, pp. 3–7.

Carman, J & Fredericks, K 2008, 'Editors' notes', in J Carman & K Fredericks (eds) *Nonprofits and evaluation: New directions for evaluation, 119*, Jossey-Bass, San Francisco.

Carman, J, Fredericks, K & Introcaso, D 2008, 'Government and accountability: paving the way for nonprofits and evaluation', in J Carman & K Fredericks (eds), *Nonprofits and evaluation: New directions for evaluation, 119*, Jossey-Bass, San Francisco.

Carter, J 1992, 'Evaluation in foundations', *American Journal of Evaluation*, vol. 13, no. 1, pp. 33–35.

Chelimsky, E 2001, 'What evaluation could do to support foundations: a framework with nine component parts', *American Journal of Evaluation*, vol. 22, no. 1, pp. 13–28.

Chin, J 2006, *The power of learning: funders and grantees getting smarter together*, Blueprint Research & Design, San Francisco, viewed 11 December 2008, <http://www.blueprintrd.com/text/power.pdf>.

Chinman, M, Imm, P & Wandersman, A 2004, *Getting to outcomes: promoting accountability through methods and tools for planning, implementation and evaluation*, Rand Corporation, Santa Monica, California.

Conner, R, Kuo, V, Melton, M & Millett, R 2004, 'Adapting evaluation to accommodate foundations' structural and cultural characteristics', in M Braverman, N Constantine & J Slater (eds), *Foundations and evaluation: contexts and practices for effective philanthropy*, Jossey-Bass, San Francisco.

Cunningham C 2003, 'Application of the TPK guidelines for conducting evaluation research with Maori', in N Lunt, C Davidson & K McKegg (eds), *Evaluating policy and practice: a New Zealand reader*, Pearson Education, Auckland.

Davidson, C 2003, 'The mix of methods: common research techniques used in evaluation', in N Lunt, C Davidson & K McKegg (eds), *Evaluating policy and practice: a New Zealand reader*, Pearson Education, Auckland.

Davidson, C, Lunt, N & McKegg, K 2003, 'Towards a conclusion: where to now for evaluation in New Zealand', in N Lunt, C Davidson & K McKegg (eds), *Evaluating policy and practice. a New Zealand reader*, Pearson Education, Auckland.

Duignan, P 2003, 'Approaches and terminology in programme and policy evaluation', in N Lunt, C Davidson & K McKegg (eds), *Evaluating policy and practice. a New Zealand reader*, Pearson Education, Auckland.

Foundation Center 2003, *Foundation yearbook: facts and figures on private and community foundations*, Foundation Center, New York.

Grantmakers for Effective Organisations 2005, *Evaluation as a pathway to learning. current topics on evaluation for grantmakers*, GEO, Washington, DC.

Grantmakers for Effective Organisations 2007, *General operating support: a GEO action guide*, GEO, Washington, DC.

Gray, S (ed.) 1993, *Leadership is—a vision of evaluation: a report of learnings from the Independent Sector's work on evaluation*, Independent Sector, Washington, DC.

Guthrie, K, Louie, J, David, T & Foster, C 2005, *The challenge of assessing policy and advocacy activities—strategies for a prospective evaluation approach: report for the California Endowment*, Blueprint Research & Design, San Francisco.

Hall, P 2004, 'A historical perspective on evaluation in foundations', in J Slater, N Constantine & M Braverman (eds.), *Foundations and evaluation*, Jossey-Bass, San Francisco.

Hamilton, R, Brown, P, Chaskin, R, Fiester, L, Richman, H, Sojourner, A & Weber, J 2005, *Learning for community change: core components of foundations that learn*, Chapin Hall discussion paper, Chapin Hall, Center for Children, Chicago.

Hendricks, M, Plantz, M, & Pritchard, K 2008, 'Measuring outcomes of United Way-funded programs: expectations and reality', in J Carman & K Fredericks (eds), *Nonprofits and evaluation,* Jossey-Bass, San Francisco.

Hernandez, G & Visher, M 2001, *Creating a culture of inquiry: changing methods—and minds—on the use of evaluation in nonprofit organizations*, The James Irvine Foundation, San Francisco, viewed 11 December 2008, <http://www.irvine.org/assets/pdf/pubs/evaluation/Creating_Culture.pdf>

Hoefer, R 2000, 'Accountability in action? Program evaluation in nonprofit human service agencies', *Nonprofit Management and Leadership*, vol. 11, no. 2, pp. 167–177.

Hoff, M 1999, 'Are 'do-gooders' really doing good?', electronic version, *Los Angeles Business Journal*, 15 November, viewed 5 January 2009, <http://www.allbusiness.com/specialty-businesses/non-profit-businesses/346295-1.html>.

Hoole, E 2007, 'Evaluation capacity building and the role of funders', paper presented at the annual meeting of the American Evaluation Association, Portland, Oregon, November.

Hoole, E & Patterson, T 2008, 'Voices from the field: evaluation as part of a learning culture', in J Carman & K Fredericks (eds), *Nonprofits and evaluations: New directions for evaluation, 119*, Jossey-Bass, San Francisco.

Irwin, K 1994, 'Maori research methods and processes: an exploration', *Sites*, vol. 28, Autumn, pp. 25–43.

Kramer, M, & Bickel, W 2004, 'Foundations and evaluation as uneasy partners in learning', in M Braverman, N Constantine & J Slater (eds), *Foundations and evaluation: contexts and practices for effective philanthropy*, Jossey-Bass, San Francisco.

Kramer, M, Graves, R, Hirschhorn, J & Fiske, L 2007, *From insight to action: new directions in foundation evaluation*, FSG Social Impact Advisors, Boston, Massachusetts.

Leviton, L & Bass, M 2004, 'Using evaluation to advance a foundation's mission', in M Braverman, N Constantine & J Slater (eds), *Foundations and evaluation: contexts and practices for effective philanthropy*, Jossey-Bass, San Francisco.

Mark, M & Beery, W 2004, 'Making judgements about what to evaluate and how intensely', in M Braverman, N Constantine & J Slater (eds), *Foundations and evaluation: contexts and practices for effective philanthropy*, Jossey-Bass, San Francisco.

McNelis, R & Bickel, W 1996, 'Building formal knowledge bases: understanding evaluation use in the foundation community', *American Journal of Evaluation,* vol. 17, no. 2, pp. 19–41.

Millett, R 2006, 'Evaluation use within foundations', unpublished paper, Millett & Associates, Chicago.

Moewaka Barnes, H 2003, 'Maori and evaluation: some issues to consider', in N Lunt, C Davidson & K McKegg (eds), *Evaluating policy and practice: a New Zealand reader*, Pearson Education, Auckland.

Moewaka Barnes, H 2009, *The evaluation hikoi: a Maori overview of programme evaluation*, Te Ropu Whariki (Whariki Research Group), Auckland, viewed 29 June 2009, <http://www.shore.ac.nz/publications/Maori%20Evaluation%20Manual%2009.pdf>.

Office of Management and Budget 2004, *Program evaluation: what constitutes strong evidence of program effectiveness?*, OMB, Washington, DC, viewed 9 December 2007, <http://www.whitehouse.gov/omb/part/2004_program_eval.pdf>.

Ostrower, F 2004, *Attitudes and practices concerning effective philanthropy*, Urban Institute Press, Washington, DC.

Patrizi, P 2006, 'The evaluation conversation: a path to impact for foundation boards and executives', *Practice matters: the improving philanthropy project*, no. 10, Foundation Center, New York, viewed 15 December 2008, <http://foundationcenter.org/gainknowledge/research/pdf/practicematters_10_paper.pdf>.

Patrizi, P & McMullan, B 1998, 'Evaluation in foundations: the unrealised potential', unpublished manuscript, WK Kellogg Foundation, Battle Creek, Michigan.

Philanthropy New Zealand 2008, *Evaluation collaboration project—building knowledge in evaluation: briefing notes*, Philanthropy New Zealand, Wellington.

Plantz, M, Greenway, M & Hendricks, M 1997, 'Outcome measurement: showing results in the non-profit sector', in K Newcomer (ed.), *Using performance measurement to improve public and nonprofit programs: New directions for program evaluation, 75*, Jossey-Bass, San Francisco.

Porter, M & Kramer, M 1999, 'Philanthropy's new agenda: creating value', *Harvard Business Review,* vol. 77, no. 6, pp. 121–131.

Preskill, H & Torres, R 1999, *Evaluative inquiry for learning in organizations*, Sage, Thousand Oaks, California.

Putnam, K 2004, *Measuring foundation performance: examples from the field*, report prepared for the California Healthcare Foundation, Oakland, California.

Sanders, J, O'Brien, M, Tennant, M, Sokolowski, S & Salamon, L 2008, *The New Zealand non-profit sector in comparative perspective*, Office for the Community and Voluntary Sector, Wellington.

Scriven, M 1994, 'Evaluation and the do-gooder fallacy', *American Evaluation Association Newsletter*, vol. 1, no. 1, pp. 6–7.

Slack, A & Leung-Wai, J 2007, *Giving New Zealand 2006*, Business and Economic Research Ltd (BERL), Wellington.

Slater, J, Constantine, N & Braverman, M 2004, 'Introduction: putting evaluation to work for foundations', in M Braverman, N Constantine & K Slater (eds), *Foundations and evaluation: contexts and practices for effective philanthropy*, Jossey-Bass, San Francisco.

Smith, L 1996, *Kaupapa Maori Research*, Department of Education, University of Auckland.

Smith, L 1999, *Decolonizing methodologies: research and indigenous peoples*, Zed Books, London.

Snibbe, A 2006, 'Drowning in data', *Stanford Social Innovation Review,* Fall, pp. 39–45.

United Way of America 1996, *Measuring program outcomes: a practical approach*, United Way of America, Alexandria, Virginia.

Urban Institute, 2004, *Key steps in outcome management*, Urban Institute, Washington, DC, viewed 18 December 2008, <http://www.urban.org/UploadedPDF/310776_KeySteps.pdf>.

Urban Institute 2010, *Quick facts about nonprofits*, National Center for Charitable Statistics, Washington, DC, viewed 9 February 2010, <http://nccs.urban.org/statistics/quickfacts.cfm>.

Urban Institute & The Center for What Works 2006, *Building a common outcome framework to measure non-profit performance*, Urban Institute, Washington, DC, and The Center for What Works, Chicago, viewed 18 December 2008, <http://www.urban.org/UploadedPDF/411404_Nonprofit_Performance.pdf>.

Wisely, D 2002, 'Parting thoughts on foundation evaluation', *American Journal of Evaluation*, vol. 23, no. 2, pp. 159–164.

WK Kellogg Foundation 1998, *WK Kellogg Foundation evaluation handbook*, WK Kellogg Foundation, Battle Creek, Michigan, viewed 18 December 2008, <http://www.wkkf.org/knowledge-center/resources/2010/W-K-Kellogg-Foundation-Evaluation-Handbook.aspx>.

Woodwell, W 2005, *Evaluation as a pathway to learning*, Current topics in evaluation series, Grantmakers for Effective Organizations, Washington, DC.

York, P 2003, *Learning as we go: making evaluation work for everyone*, briefing paper, TCC Group, New York, viewed 11 December 2008, <http://www.tccgrp.com/pdfs/per_brief_lawg.pdf>.