

BUILDING SOCIAL POLICY EVALUATION CAPACITY

Paul Duignan¹
Senior Lecturer
Alcohol & Public Health Research Unit
University of Auckland

Abstract

The last three years have seen an increasing interest in evaluation in the public sector in New Zealand. This trend could result in an adequately resourced and sophisticated approach to evaluation, involving policy and provider levels within government, Māori and third-sector/community organisations. This in turn could lead to better formed and implemented social programmes and policies. On the other hand, it is possible that unrealistic expectations, an unsophisticated model of evaluation, lack of strategic involvement of stakeholders and inadequate investment in appropriate evaluation capacity building will result in the current wave of enthusiasm ultimately turning to disillusionment. If we use the current increased interest in evaluation to build and embed a sophisticated evaluation capacity across the social policy sector we are likely to see a more positive outcome. To achieve this we need to use appropriate evaluation models, including those appropriate for Māori programmes; build a sector culture of evaluation through appropriate evaluation training and awareness-raising at all levels; and attempt to foster strategic, sector-wide priority setting of evaluation questions.

INTRODUCTION

The final years of the last decade saw mounting interest in evaluation and an outcomes focus within the New Zealand social policy community (Schick 1996, Bushnell 1998, Duignan 1999, State Services Commission 1999, Controller and Auditor-General 2000). From the point of view of the working evaluator, this seems to have been accompanied by a significant

¹ **Acknowledgements**

The author would like to acknowledge all of those who have worked at the Alcohol & Public Health Research Unit (APHRU) and the Māori research unit Whariki which works in partnership with APHRU. He would also like to acknowledge Jane McCann and Carolyn Lane of the organizational development company Navigate where he worked while undertaking additional social sector evaluation work.

This paper draws in part on the thinking behind "Mainstreaming Evaluation or Building Evaluation Capability?" (Duignan 2001), which won the American Evaluation Association's President's Prize.

Correspondence

Dr. Paul Duignan, APHRU, University of Auckland (www.aphru.ac.nz, p.duignan@auckland.ac.nz); Parker Duignan Ltd (www.parkerduignan.com, paul@parkerduignan.com)

rise in the amount of evaluation being funded and undertaken in New Zealand. It will be fascinating to watch how this develops over the next decade. If we are lucky it will result in more sophisticated evaluation being undertaken, which will feed into the formation and implementation of better social policy. If we are unlucky there is likely to be an initial burst of evaluation activity for a few years with a lot of resources spent on elaborate technical evaluation designs, followed by a phase of disillusionment due to unrealistic expectations as to what evaluation can deliver for social policy in New Zealand.

If we are to get the most out of the increased interest in evaluation we must build an enduring evaluation capacity in the social policy area. Part of this involves increasing the number of evaluators involved with the sector, as has been done in some evaluation capacity building (Compton et al. 2001), but it needs to go beyond this to put in place the following three elements:

- using appropriate evaluation models;
- developing a culture of evaluation throughout the social policy sector by teaching evaluation skills appropriate for each level of the sector; and
- sector-level strategising to identify priority evaluation questions, rather than just relying on evaluation planning at the individual programme level.

Each of these needs to involve government, community organisations and Māori stakeholders in the development of a more strategic approach to social policy evaluation.

USING AN APPROPRIATE EVALUATION MODEL

Discussing an appropriate evaluation model may seem a slightly obscure and theoretical place to start thinking about building social policy evaluation capacity. However, there are a number of different ways in which evaluation can be described, and various models and typologies that are in use by evaluators (Cook and Campbell 1979, McClintock 1986, Patton 1986, Guba and Lincoln 1989, Rossi and Freeman 1989, Scriven 1991, Fetterman et al. 1996, Chelimsky and Shadish 1997). From the author's experience, these models and approaches are not all the same in terms of their suitability for social policy evaluation capacity building. Suitable evaluation models should:

- attempt to demystify evaluation so that it can be understood and practised at all levels within the social policy sector;
- use a set of evaluation terms that emphasises that evaluation can take place across a programme's life cycle and is not limited to outcome evaluation;
- allow a role for both internal and external evaluators;
- have methods for hard-to-evaluate, real-world programmes, not just ideal-type, large-scale, expensive, external evaluation designs;
- not privilege any one meta-approach to evaluation (for example, goal-free, empowerment);

- be based on a sophisticated understanding of what evaluation can actually deliver in terms of an evidence base for social policy; and
- take into account the need for approaches for evaluating Māori programmes that may be different from mainstream evaluation approaches.

Some evaluation models meet these criteria better than others. Each of the criteria is discussed below.

Demystifying Evaluation

An appropriate evaluation model for social policy evaluation capability building should be able to be explained in clear terms to a wide range of different stakeholders with diverse training, backgrounds and experience from across government, Māori and the community sectors. Such a model must at the same time be able to accommodate complex technical evaluation methodologies within this easily understandable framework.

One way to describe evaluation for capacity building is to conceptualise it as being about asking questions – of our programmes, organisations and policies. These questions are not something that evaluators alone should attempt to answer themselves; they are questions that should be an important concern of every policy maker, manager, staff member and programme participant. The high-level question I use in describing evaluation is always:

- Is this (organisational activity, policy or programme) being done in the best possible way?

This is then unpacked into a series of subsidiary questions:

- How can we improve this organisation, programme or policy?
- Can we describe what is happening in this organisation, programme or policy?
- What have been the intended or unintended outcomes from this organisation, programme or policy?

A question-based introduction to evaluation helps to demystify the process of evaluation. It puts the responsibility for evaluation back where it belongs – on the policy makers, funders, managers, staff and programme participants to identify the questions they are interested in, rather than leaving it solely with evaluators. It highlights that programme managers and staff cannot avoid these questions; they just have to work out ways of answering them. In most cases stakeholders will have to answer these questions through their own efforts. However, in some instances they will need to call in specialised evaluation assistance. A question-based approach to evaluation is also well positioned to highlight the concept of sector-level strategising about priority evaluation questions, which is discussed later in this article.

A Set of Evaluation Terms That Apply Across the Programme Life Cycle

In New Zealand, at least, most stakeholders unfamiliar with evaluation still see it mainly in terms of outcome evaluation, although this narrow perspective is now starting to change. An appropriate set of terms for the different types of evaluation should highlight that evaluation consists of much more than this. Two important dichotomies are often used to describe evaluation: the distinction between *formative* and *summative* evaluation and the distinction between *process* and *outcome* evaluation. Combining elements from both leads us to a three-way typology – formative, process and impact/outcome – that emphasises that evaluation can take place right across the programme life cycle, not just at the end. This is the three-way split used in the evaluation work of the Alcohol & Public Health Research Unit (Casswell and Duignan 1989, Duignan 1990, Duignan and Casswell 1990, Duignan et al. 1992a, Duignan et al. 1992b, Turner et al. 1992, Duignan 1997, Waa et al. 1998, Casswell 1999, Health Research Council n.d.).

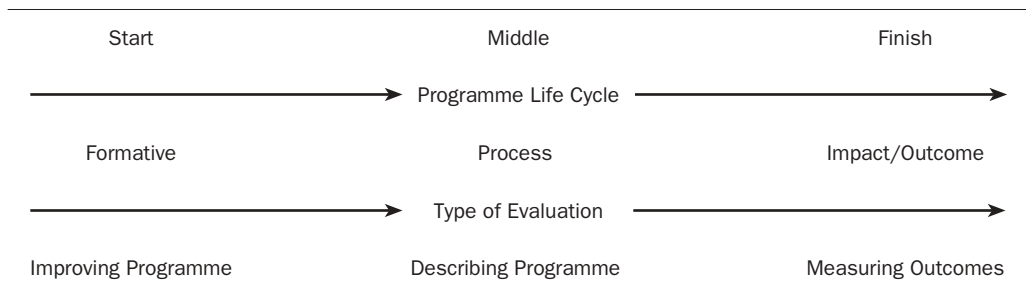
In this typology, which is based on the purpose for which evaluation will be used, formative evaluation (McClintock 1986, Dehar et al. 1993, Tessmer 1993) is defined as *evaluation activity directed at optimising a programme*. (It can, alternatively, be described as design, developmental or implementation evaluation).

Process evaluation (Scheirer 1994) is defined in our typology as *describing and documenting what happens in the context and course of a programme to assist in understanding a programme and interpreting programme outcomes, and/or to allow others to replicate the programme in the future*. Note that this narrows the definition of process evaluation by not including the formative evaluation element.

Outcome evaluation (Cook and Campbell 1979) is defined in the typology as *assessing the positive and negative results of a programme*. This includes all sorts of impact/outcome measurement, recognising that outcomes can be short, intermediate or long term and also arranged in structured hierarchies (for example, individual level, community level, policy level).

None of these terms are opposed to each other. They are seen as three essential purposes for evaluation. The three terms can in turn be directly related to the three subsidiary evaluation questions identified in the section above. They can also be related to the start, middle and end of a programme. This encourages thinking about how evaluation can be used right across a programme's life cycle. Each type of evaluation – formative, process and impact/outcome – must be individually considered as a possibility for evaluation activity. If outcome evaluation proves too expensive or difficult, there may still be useful questions that can be answered about formative and process evaluation.

Figure 1 The Relationship Between Types of Evaluation and Stages in the Programme Life Cycle



Internal and External Evaluators

An appropriate evaluation model for building evaluation capability must also allow for the possibility of both internal and external evaluators (Mathison 1991, Minnett 1999). If evaluation is seen as something that is only undertaken by external experts then there is little reason for internal staff to improve their evaluation skills. This is particularly relevant for Māori and community-sector organisations, often with little access to outside evaluation resources. A useful evaluation model for capacity building needs to have plenty to offer the internal evaluator with limited resources for evaluation, rather than just focusing on the needs and concerns of the relatively well-resourced external evaluator. It is more likely that formative and process evaluation techniques will be the ones that are possible within the usually limited resources available to internal evaluators.

Methods for Hard-to-Evaluate Real-World Programmes

An appropriate evaluation model for capability building also needs to incorporate methods that can be used to evaluate a wide range of real-world programmes that tend to present interesting evaluation challenges. One area where appropriate evaluation models are crucially important is in community programmes. Evaluating community-based programmes presents interesting challenges for evaluators and raises considerable technical and political issues for traditional models of evaluation (Edelman 2000). Community programmes have long time frames, and take place in communities where many other programmes are running at the same time, often with the same goals. Even more challenging, community programmes are usually based on a philosophy of community autonomy (Shirley 1982). This presents interesting tensions for evaluation when looking at whether or not a programme has met its objectives. Should the evaluation assess achievement of a set of objectives prescribed by the funder, or a set of objectives set by the community itself, or both? There are models and approaches that can be used in the evaluation of such programmes (Duignan and Casswell 1989, Duignan and Casswell 1992, Duignan et al. 1993, Moewaka Barnes 2000a). These

models and approaches need to be further refined as part of the essential toolkit building social policy evaluation of real-world programmes.

Not Privileging Any One Meta-Approach to Evaluation

Meta-approaches to evaluation are evaluation styles that endorse a particular solution to the philosophy of science questions underlying evaluation – in particular stakeholders interest in the truth status of claims that are made in an evaluation. Goal-free evaluation (Scriven 1972) and empowerment evaluation (Fetterman et al. 1996) are good examples of meta-approaches to evaluation that take different philosophy of science positions (Scriven and Kramer 1994). It is fine for evaluators to adopt one or other of these meta-positions in their professional work as evaluators. However, in building evaluation capability it is important that a more inclusive approach is taken to evaluation that does not privilege just one approach. Of course, the Western evaluation approach itself can be seen as just one meta-approach to evaluation and we need to be aware that this is not universally accepted by stakeholders. Māori are actively involved in the process of developing evaluation models and approaches that may or may not have similar assumptions, methods, and techniques to evaluation as it is practised in the Western tradition (Watene-Heydon et al. 1995, Moewaka Barnes 2000a, Smith 2000, Moewaka Barnes 2000b).

A Sophisticated Model of the Evidential Base That Evaluation Can Deliver

The last element in the evaluation model needed for social policy capacity building is a sophisticated model of the evidential base that evaluation is likely to be able to deliver. There is a tendency in social policy to start with a naïve expectation that evaluation may be able to deliver the type of “evidential map” that is illustrated in Figure 2.

Figure 2 Evidential Map of Links between Social Policy Programmes or Policies and Outcomes

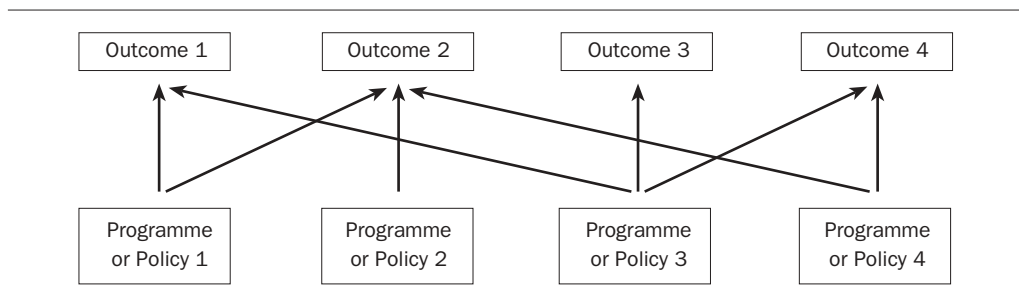
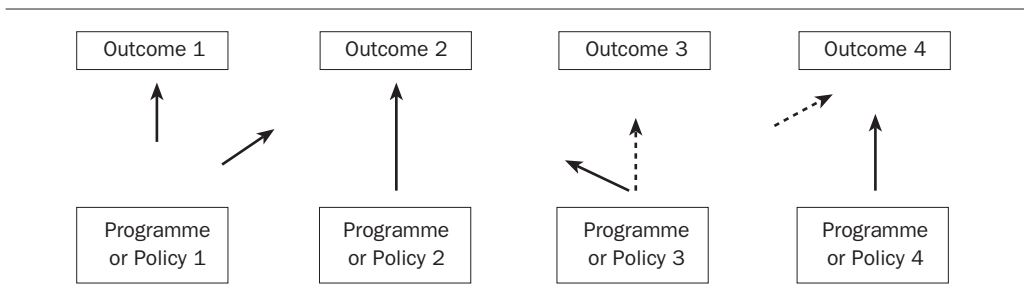


Figure 2 shows evaluation providing evidence linking a series of social policy programmes or policies to a series of cross-sector social outcomes. Everyone would acknowledge that because of resource and technical constraints, evaluation cannot provide a totally comprehensive map of these links. However, it is important to distinguish between holding the view that we can approach a comprehensive evidential map, as in Figure 2 or whether our expectations of evaluation should be much more like what is set out in Figure 3.

Figure 3 The Likely Extent of the Evidential Map Delivered Through Evaluation



In the author's view, Figure 3 provides a much more realistic picture of what evaluation is likely to be able to deliver in the social policy area, even when a large amount of evaluation is being undertaken. We are unlikely to ever get anything like a full evidential map on which to base rational social policy. We will continue to be forced to make substantial decisions under uncertainty. Within the evidential map there will, of course, be some connections between programmes and outcomes that are easier to measure than others. These relatively easy-to-evaluate programmes will tend to:

- operate at only the *individual* level rather than including *organisational*, *community* and *policy*-level strategies;
- take place in only one locality rather than at the national and local level;
- focus on single-outcome variables that are already routinely collected, rather than *multiple*-outcome variables;
- take place in institutionalised controlled settings; and
- seek outcomes that can be measured within a relatively short timeframe.

For instance, a school-based programme that uses examination results as its outcome measures is a good example of where it is relatively easy to measure and attribute changes in outcomes to the effects of a programme.

It is important that, as we increase the evaluation activity taking place in New Zealand, we are realistic about what can be provided in terms of the social policy evidential map. We also need to understand the implications for social policy decision making of certain outcome

evaluation designs being easier to implement in some social policy areas than in others. We cannot afford to become too simplistic about the automatic application of outcome evaluation results to determining priorities for funding programmes and policies. This is particularly important as we move toward building the information base for “evidence-based” practice in social policy (Wright 1999). The fact that quasi-experimental outcome evaluations are possible in some social policy areas should not be taken as evidence that *quasi-experimental outcome evaluation is similarly feasible in all policy areas*, where alternatives such as case study designs may need to be used. The amount of experimental outcome evidence for different types of programmes and policies is a function of both the actual effectiveness of the programmes and the *ease of undertaking experimental outcome evaluations on the type of programme under consideration*.

Given the current interest in “joined-up solutions” in the social policy area (Maharey 2000), many of the programmes and policies currently being proposed have characteristics that mean they are more difficult to evaluate. They tend to:

- use a range of strategies at the individual, community and policy level in an integrated programme;
- take place at both the local and national level at the same time;
- be directed at multiple rather than single outcomes, some of which may be expensive to collect data on;
- take place in uncontrolled community, rather than institutional, settings; and
- seek long-term outcomes that will take years to come to fruition.

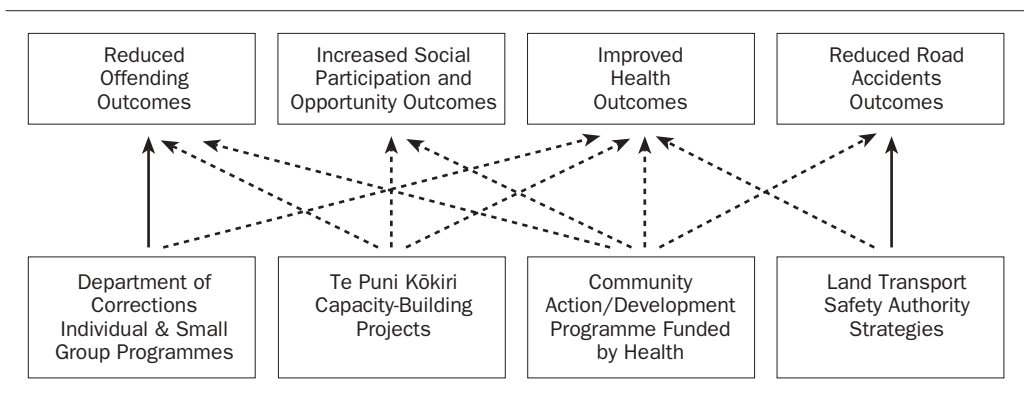
In these cases experimental outcome evaluation is much more difficult. This does not mean that we should not attempt to undertake evaluations of such programmes, but that the evaluation designs we use will have to be different. These evaluation designs, such as case studies, will yield different types of data from the quasi-experimental designs. A more sophisticated approach needs to be taken to evaluating such programmes, using a range of types and methods of evaluation; as discussed early in this paper. There will of course still be situations in which experimental or quasi-experimental outcome evaluation is possible and should be undertaken if it will answer a priority evaluation question for the sector.

The issue of how comprehensive an evidential map evaluation can provide becomes particularly critical when attempting an evidence-based approach to prioritising interventions to achieve cross-sector social policy objectives. The author’s experience during a recent review of strategic social policy for the Ministry of Social Policy and the State Services Commission indicated that this sort of prioritisation was, naturally enough, on the wish list of politicians and policy analysts alike (Duignan and Stephens 2001). However, such exercises can never become routinely empirically based (at least for the foreseeable future). illustrates this point by looking at what the evidential map may look like in a limited selection of cross-linked social programme areas.

Looking first at the left-hand side of Figure 4, it may be possible to use outcome evaluation results to attribute the outcome of a reduction in offending to small group programmes run by the Department of Corrections. However it is also likely that capacity building programmes, funded by Te Puni Kōkiri, working at the community level (and hence harder to evaluate for attribution of outcomes) may be also be contributing to reduced offending. Equally, a community action programme funded through the Health Vote that includes an element of reducing alcohol abuse may also reduce offending.

Another example is provided on the right-hand side of the figure, where a reduction in road accidents can be directly attributed through evaluation to Land Transport Safety Authority activity. However, a community action programme from the Health Vote may also reduce road accidents due to its targeting of alcohol abuse (but again, because it is a community-based strategy it is likely to be more difficult to provide outcome evaluation evidence to directly attribute reduced road accidents at the same level of certainty as can be done for Land Transport Safety Authority activity).

Figure 4 An Example of the Evidential Map for a Selection of Social Programmes



EVALUATION MODELS FOR EVALUATING PROGRAMMES FOR MĀORI

An additional element, which is essential to any evaluation model used in capacity building in the social policy area, is providing for the evaluation of programmes for Māori. Since evaluation is about placing value on policies and programmes, it raises enormous issues in New Zealand in the context of the Treaty of Waitangi. A number of priority government programmes are currently targeted at Māori, and evaluation of these programmes obviously needs to be undertaken. Various guidelines have been issued regarding research and Māori (Te Puni Kōkiri 1999, Health Research Council n.d.). The most basic consideration is whether methods used in an evaluation are going to work with Māori.

However, the issue of evaluating programmes for Māori is much wider than this. There are considerations at the level of the evaluation's:

- governance and control;
- cultural integrity;
- stakeholder support;
- personnel;
- methods used; and
- dissemination and control of data.

While there is considerable ongoing discussion of research methods from a Māori perspective (Smith 2000), specific models for the evaluation of Māori programmes are still emerging. In the longer term, a project currently under way, funded by Te Puni Kōkiri, is looking separately at government expectations of evaluation of Māori programmes and Māori perspectives on evaluation models, looking for commonalities and divergences in these perspectives. Out of this, recommendations may emerge regarding support for current and future work on developing models that work for Māori.

BUILDING A CULTURE OF EVALUATION

The next step in building evaluation capability is to develop skills, systems and structures for evaluation activity at all levels within the social policy sector, government, community sector and Māori. The objective of skills development is to make evaluation more sophisticated (along the lines of the evaluation model discussed above), and to teach appropriate specific evaluation skills to those who can use them in their day-to-day work. This can be done by developing appropriate training materials and by running training workshops.

The Alcohol & Public Health Research Unit and Whariki have developed a series of manuals on evaluation that reflect the evaluation model described above, and these have been widely distributed throughout the public and community health sector in New Zealand (Casswell and Duignan 1989, Duignan et al. 1992b, Turner et al. 1992, Waa et al. 1998). The response to these manuals has been evaluated and subsequent manuals have been amended on the basis of this feedback.

During the time the resources have been available, the Unit and Whariki have carried out a series of different types of training workshops for different audiences within the sector.

- Brief presentations on evaluation can be included in a range of sector workshops on other issues. Typically these are one to two hours long, and cover the general evaluation model and principles, raising awareness of evaluation within the sector.
- Two-day Level I (first-year tertiary) courses for service-provider lower-level managers and staff provide them with an opportunity to discuss the overall evaluation model and learn

specific evaluation skills they can use in their day-to-day work. Considerable time is spent demystifying evaluation and describing simple formative and process evaluation methods that can be used by service provider staff. Suitable outcome evaluation methods are discussed, as are indicators of when other evaluation expertise needs to be drawn in.

- More advanced Level II two-day courses are aimed at service-provider managers and staff wanting to develop their skills. These provide more in-depth training in evaluation skills.
- One-week-long workshops allow for policy makers, funders, larger provider specialists, and researchers to develop and practise appropriate evaluation skills. These cover the evaluation model and the skills and techniques discussed in the Level I and II workshops in more depth, with further discussion of outcome evaluation issues.
- Workshops are run by Whariki Māori evaluators specifically for Māori programme management and staff. These look at evaluation concepts and methods from a Māori perspective.
- One-day overview workshops discuss evaluation concepts and approaches for service-provider management. These discuss the concepts from the evaluation model and how they relate to organisational policies and practices.
- One-day workshops for staff and management within an organisation cover the evaluation model, concepts in evaluation and the idea of prioritising evaluation questions across the organisation as a whole.
- Postgraduate university masters' papers are given at the University of Auckland for researchers and practitioners interested in further developing their understanding of evaluation and their ability to undertake evaluations.

All of these courses, apart from those for managers, involve discussion of evaluation models combined with hands-on working with evaluation projects brought to the workshops by participants. This action learning approach ensures that participants go away with a feeling of mastery in at least some evaluation techniques, which further helps to promote the idea that there are aspects of evaluation that can be done by people at all levels within a programme, organisation or sector.

There are signs that this sort of approach is starting to be utilised in other parts of the social policy sector. A number of government agencies have been building their evaluation units and obtaining evaluation training for their staff. Under the auspices of the Australasian Evaluation Society evaluation courses have been run in Wellington as part of this. The most systematic approach that has come to my attention is the capacity-building work developed by Department of Work and Income (now part of the Ministry of Social Development) within its own organisation to boost its evaluation capacity. Given the Ministry's key role in social policy it will be interesting to see if such capacity-building approaches spread to other parts of the social policy sector.

ORGANISATIONAL OR SECTOR-LEVEL STRATEGISING TO PRIORITISE EVALUATION QUESTIONS

The last essential element for social policy sector evaluation capacity building is to foster sector-level strategising about the priority evaluation questions for developing social policy in New Zealand. These discussions need to involve government, Māori and community-sector stakeholders.

In thinking about such strategic discussions it is important to distinguish discussions about *evaluation* from discussions about *performance measurement*. Performance measurement is a routine activity that uses relatively cheap and easily collected measures to monitor whether a programme is delivered. Evaluation, on the other hand, is generally a more strategic, more expensive, more detailed one-off activity. This can include the usually more difficult-to-measure question of whether the programme is achieving its objectives. Both performance measurement and evaluation are essential. However, when their purposes become confused and one starts to predominate, as some commentators believe has become the case in North America, this causes major problems (Perrin 1998). One mistake is inappropriate attempts to make unrealistic evaluation-type evidential demands on a programme to “prove” that it is achieving hard-to-measure outcomes as part of its routine performance measurement.

The relationship between strategy, evaluation and performance measurement can be characterised as set out in Figure 5.

Figure 5 Relationship between Strategy, Evaluation and Performance Measurement

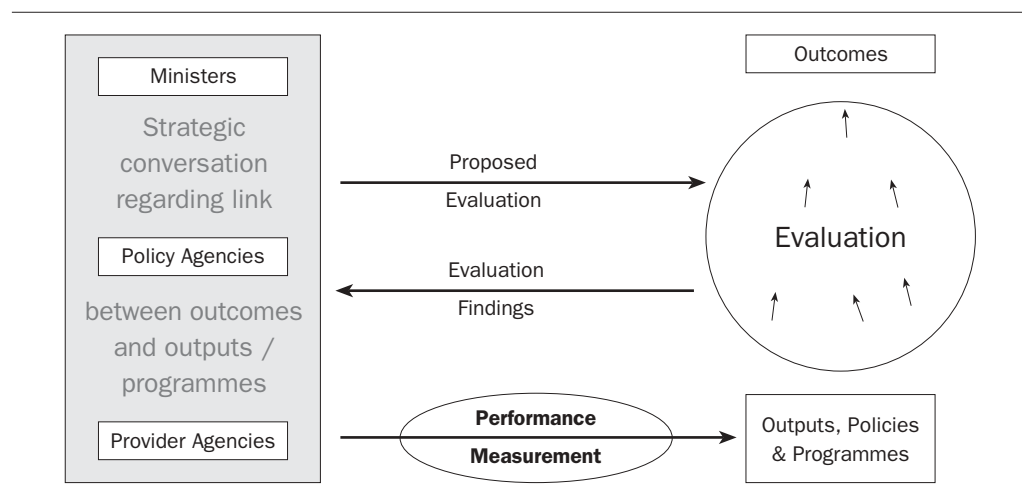


Figure 5 is based on the notion that decision making about outputs, performance management and evaluation should take place as a part of a strategic discussion between Ministers, central agencies and other agencies. These strategic discussions should involve Māori, government and community-sector agencies in the roles of both policy player and provider. These discussions, which will take place in a number of forums, should:

- take into account all the evaluation findings that have been established up to that time to determine what outputs/programmes are likely to achieve the outcomes being sought by the government;
- specify the outputs/programmes that need to be delivered by agencies in the next period;
- specify performance management processes to ensure that the agencies deliver the outputs/programmes they have agreed to deliver; and
- determine the next round of *priority evaluation questions* that will better inform the next stage of the strategic conversation.

The result of this sort of approach should be to move the emphasis away from thinking about evaluation as a “technical” programme towards seeing evaluation as contributing to the ongoing strategic social policy debate. A less technically sophisticated evaluation that can add more to the strategic conversation is of more value than a technically sophisticated evaluation whose results may distract from ensuring the most strategic use of scarce public sector resources.

One useful way of looking at evaluation from this different perspective is to see the spending on evaluation as part of one big research and development fund available for social policy development. This highlights the role of the government, central agencies, Māori and the community sector in defining evaluation questions, rather than seeing evaluation as something that is “done to” programmes or agencies. This also helps to move away from some of the undesirable aspects of the growth of an “evaluation industry” based around individual programmes and agencies which tend to stem from obligatory programme and agency evaluation regimes (Bushnell 1998).

Of course, it can be argued that already a lot of organisational and sector strategic considerations are factored into the evaluation requirements for an individual programme. Funders will indicate which programmes they want evaluated and the level of resources, and may indicate which evaluation questions they want answered. In addition, in reviews of the academic literature, and in priority-setting processes within research-funding bodies, there prioritisation will be happening. In those sectors where there are ongoing research groups involved in teaching, advising and undertaking a large number of evaluations, they will play a role through having a strategic view of a sector and the evaluation questions that are to be the next priority. The Alcohol and Public Health Research Unit and the Injury Prevention Unit at the University of Auckland have played this role for their respective sectors over a number of years. However,

for the social policy sector this process is still too ad hoc and there is often a disjunction between priority setting (where it is taking place) and what actually happens on the ground with the evaluation of the many programmes that are subject to evaluation.

Exactly how to facilitate setting evaluation priorities in the social policy sector, outside of the contribution made by dedicated research units, is a difficult question. Sectors dealing with social issues tend to be made up of a diverse range of public and private groups funding a diverse range of programmes. There are some innovative evaluation priority-setting exercises going on in New Zealand at the moment in the labour and employment programme area (Kate McKegg personal communication).

CONCLUSION

This paper has looked at developing evaluation capacity in the social policy sector in New Zealand. The increased interest in evaluation has been greeted enthusiastically by evaluators who have been working in evaluation in New Zealand over the last few decades. It is particularly welcome after the lack of adequate attention to outcomes that occurred following implementation of the outputs/outcomes dichotomy introduced at the time of the public sector reforms in the 1980s.

If we are to make the most of the current interest in evaluation, we need to focus on building evaluation capacity at all levels and in all parts of the social policy sector. This will require a rather different approach from solely developing large-scale technically sophisticated evaluations of social policy interventions. There is a place for such evaluations when they are targeted at answering strategic evaluation questions for the sector, adequate resources are available, and they are feasible in terms of the programme or policy being evaluated. However, these types of evaluations need to be embedded within a social policy sector in which everybody sees evaluation as their responsibility.

Everybody within the sector also needs to be equipped with the skills to undertake appropriate types of evaluation in their day-to-day work. What we need now are resources for training and awareness-raising in regard to appropriate evaluation models and methods, and hard thinking about the difficult task of how we can foster richer and more comprehensive discussions to identify the priority evaluation questions across the social policy sector.

REFERENCES

- Bushnell, P. (1998) "Does evaluation of policies matter?" *Evaluation*, 4(3):363-371.
- Casswell, S. (1999) "Evaluation research" in C. Davidson and M. Tolich (eds.) *Social Science Research in New Zealand: Many Paths to Understanding*, Longman, Auckland.
- Casswell, S. and P. Duignan (1989) *Evaluating Health Promotion: A Guide for Health Promoters and Health Managers*, Department of Community Health, School of Medicine, University of Auckland.
- Chelimsky, E. and W. R. Shadish (eds.) (1997) *Evaluation for the 21st century: a handbook*, Sage, Thousand Oaks, California.
- Controller and Auditor-General (2000) *First Report for 2000: Health, School Board of Trustees, Impact Evaluation*, New Zealand Office of the Controller and Auditor-General, Wellington.
- Cook, T. and D. T. Campbell (1979) *Quasi-Experimentation: Design and Analysis Issues for Field Settings*, Houghton Mifflin Company, Boston.
- Duignan, P. (1990) "Evaluating health promotion: an integrated framework" presented at "Health Promotion Research Methods: Expanding the Repertoire" Conference, Toronto, Canada.
- Duignan, P. (1997) *Evaluating health promotion: the Strategic Evaluation Framework*, D.Phil., Community Psychology, University of Waikato, Hamilton.
- Duignan, P. (1999) "Summary of 'Improving Public Sector Policy Through Quality Evaluation Conference', Wellington, New Zealand, 25-28 May 1999" [http://www.navigate.co.nz/discuss.htm#Evaluation Articles](http://www.navigate.co.nz/discuss.htm#Evaluation%20Articles).
- Duignan, P. (2001) "Mainstreaming evaluation or building evaluation capability?" presented at "Evaluation 2001" conference of The American Evaluation Association, St. Louis, Missouri, November 7-10.
- Duignan, P. and S. Casswell (1989) "Evaluating community development programs for health promotion: problems illustrated by a New Zealand example" *Community Health Studies*, 13(1): 74-81.
- Duignan, P. and S. Casswell (1990) "Appropriate evaluation methodology for health promotion" presented at American Evaluation Society Annual Conference, Washington, D.C.
- Duignan, P. and S. Casswell (1992) "Community alcohol action programme evaluation in New Zealand" *Journal of Drug Issues*, 22:757-771.
- Duignan, P. and B. Stephens (2001) *Review of the Ministry of Social Policy's Strategic Social Policy Capability*, State Services Commission and Ministry of Social Policy, Wellington.
- Edelman, I. (2000) "Evaluation and community-based initiatives" *Social Policy*, 31(2):13-23.
- Guba, E.G. and Y.S. Lincoln (1989) *Fourth Generation Evaluation*, Newbury Park, California, Sage.
- Health Research Council (n.d.) "Guidelines for Researchers on Health Research Involving

- Māori” <http://www.hrc.govt.nz/Maoguide.htm#Introduction>.
- Maharey, S. (2000) “Social Directions Briefing” <http://www.executive.govt.nz/minister/maharey/directions/index.html>.
- Mathison, S. (1991) “Role conflict for internal evaluation” *Evaluation and Program Planning*, 14:173-179.
- McClintock, C. (1986) “Towards a theory of formative program evaluation” *Evaluation Studies Review Annual*, 11:205-223.
- Minnett, A. M. (1999) “Internal evaluation in a self-reflective organization: one nonprofit agency’s model” *Evaluation and Program Planning*, 22(3):353-362.
- Moewaka Barnes, H. (2000a) “Collaboration in community action, a successful partnership between indigenous communities and researchers” *Health Promotion International*, 15:17-25.
- Moewaka Barnes, H. (2000b) “Kaupapa Māori: explaining the ordinary” *Pacific Health Dialog*, 7(1):13-16.
- Patton, M. Q. (1986) *Utilization Focused Evaluation*, Sage, Newbury Park, California.
- Perrin, B. (1998) “Effective Use and Misuse of Performance Measurement” *American Journal of Evaluation*, 19(3):367-379.
- Rossi, P. H. and H. E. Freeman (1989) *Evaluation: A Systematic Approach*, Sage, Beverly Hills, California.
- Scheirer, M. A. (1994) “Designing and using process evaluation” in J. Wholey, H. Hatry and K. Newcomer (eds.) *Handbook of Practical Program Evaluation*, Jossey-Bass, San Francisco.
- Schick, A. (1996) *The Spirit of Reform: Managing New Zealand’s State Sector in a Time of Change*, State Services Commission, Wellington.
- Scriven, M. (1972) “Pros and cons about goal-free evaluation” *Evaluation Comment: The Journal of Educational Evaluation*, 3(4):1-7.
- Scriven, M. (1991) *Evaluation Thesaurus*, Sage, Newbury Park, California.
- Scriven, M. and J. Kramer (1994) “Risks, rights and responsibilities in evaluation” *Evaluation Journal of Australasia*, 9(2):3-16.
- Shirley, I. (1982) *Development Tracks: The Theory and Practice of Community Development*, Dunmore Press, Palmerston North.
- Smith, L. T. (2000) *Decolonising Methodology: Research and Indigenous Peoples*, Zed Books, London.
- State Services Commission (1999) *Essential Ingredients: Improving the Quality of Policy Advice*, State Services Commission, Wellington.
- Te Puni Kōkiri (1999) *Evaluation for Māori: Guidelines for Government Agencies*, Te Puni Kōkiri, Wellington.
- Tessmer, M. (1993) *Planning and Conducting Formative Evaluations*, Kogan Page, London.
- Wright, J. C. (1999) *Strategic Social Policy Advice: Improving the Information Base*, Working Paper No. 8, State Services Commission.