

Evidence and Policy Making: The Evaluator's Contribution

Dr Angela O'Brien-Malone

Abstract

The use of evidence, particularly evaluation evidence, in policy-making is seen as highly desirable — indeed, it is seen as 'best practice'. However, the reality of the relationship between policy making and evidence, or perhaps to put it better, between policy-making and evidence-making, is much more complex. One of the issues that contribute to the complex nature of the relationship is policy makers' need to rely on more than evidence when making policy. Another is that policy makers need to be able to assess evidence. The UK House of Commons Science and Technology Committee (HoCSTC) recently addressed some of these problems, as has recent work from the OECD's Centre for Educational Research and Innovation (CERI). The growing professionalism of evaluation is considered in the context of these issues, together with the contribution which evaluators can make as knowledge brokers. These questions are discussed in this paper.

The last few months have seen the baby language debate getting some media coverage. The question under debate is: Does a baby learn language well from the baby DVDs now flooding the market? "No", says some research (Zimmerman, Christakis, & Meltzoff, 2007). But baby DVDs are big business — the Baby Einstein set of DVDs alone is expected to earn \$1 billion for Disney by 2010 — that is *very* big business, and that alone is enough to make it relevant to policy. In addition, the question of baby language learning and baby DVDs is also relevant to policy about early childhood. So the efficacy of baby DVDs is an issue which might be incredibly important to policy makers. And where is it being covered? It is being covered in journals like *Nature* or *American Scientist* or the *Journal of Experimental Psychology* — journals which very few policy makers have either the time or the background to read and understand.

The history of the last few years is littered with examples where policy makers, and decision makers (and here that generally means politicians), have needed to understand and assess information which goes well beyond what the average policy maker can reasonably be expected to be familiar with. These examples include: the outbreaks of SARS, foot and mouth disease and CJD; climate change; planning for an avian influenza pandemic; managing water resources; balkanisation in Europe; terrorism; persistent poverty; illiteracy; innumeracy; the uptake of fallout from Chernobyl into the food chain — the list goes on. Some of the policies and decisions which have to be made in response to such issues are extremely time-sensitive — the SARS outbreak required an *immediate* response — others, such as illiteracy and poverty, are such long-term problems that they extend well beyond the likely life of any government and impose quite different problems related to time scale of response.

What is the state of evidence about these various important policy issues? In a perfect world all the necessary evidence would always be available, but the reality is that sometimes there is little evidence available, and what there is may be lacking or incomplete. Alternatively, a relevant body of evidence may actually exist, but the bridge between the researchers who have the evidence and those who devise the policies has not been built; policy makers may simply not know of its existence. Other problems are that the available evidence may not suggest a single course of action, the evidence may be of variable quality, or the research may have yielded several apparently irreconcilable outcomes. How do policy makers manage in this messy environment? And is there a role and an opportunity for evaluators?

The points I want to spend some time on today fall under three headings:

1. the complex relationship between evidence and policy making,
2. the issue of expertise and 'knowledge-in-depth', and
3. the contribution of evaluators.

The relationship between evidence and policy making

Evidence-based policy making is often seen as 'best practice' in policy making. I certainly would not want to discourage the use of evidence in making policy! However, the reality of the relationship between policy making and evidence, or perhaps to put it better, between policy-making and evidence-making, is much more complex than it at first appears.

When Labour came to power in the UK in 1997, evaluators were delighted that that government made an early and very clear commitment to evidence-based policy making. Famously, this was summed up by Tony Blair when he declared 'what matters is what works'. What resulted from this was a cultural shift on evaluation and on the use of evidence. What has been learned since then about evidence-based policy making?

The first thing to note is that there has been a subtle but important shift in language, with many key players now preferring to talk about evidence-*informed* policy, rather than evidence-*based* policy. What does this shift signify and why is it important? This shift in language has developed in response to a growing sophistication in understanding the relationship between evidence and policy. Evidence is not the only thing which must be weighed in when developing policy. For instance, Governments are selected by voters based, in part, upon differences in ideology, so ideology must also have a legitimate role in policy making. Need also drives policy: The evidence might not be available, but policies and decisions must be made anyway. The community sometimes will not accept what the evidence suggests, and so a different tack must be taken. Sometimes money is not available, or time is too short, or the supply of skills or expertise required for a particular form of action is lacking. Sometimes the evidence clearly indicates that a particular course of action would be better — but the cost-benefit analysis shows that the after-cost gain is minimal and so, because cost must be factored in, that course of action is not pursued.

Of course, evaluators are part of the evidence-making side of the equation, but we are often closer to the policy makers, and often more aware of the policy consequences, than other contributors to the evidence-making side of the equation. So our role as evaluators produces opportunities and, possibly, responsibilities that others engaged in the evidence-making side of the equation may not have.

Knowledge-in-depth and what can be expected of policy makers

When corporations talk about hostile takeovers, or designers of nuclear power plants consider the possibility of failure, a phrase which comes up is 'defence-in-depth' — this relates to the idea of a structure of layer after layer of defensive measures. It's a good phrase and it captures well the broad outline of the strategy. In an equivalent sort of way, I want to introduce the notion of knowledge-in-depth. I want to distinguish this from what is commonly called in-depth knowledge. Having in-depth knowledge is having a great deal of knowledge about a narrow domain of human knowledge. What I mean by knowledge-in-depth is something much broader, it is knowledge of the layer after layer of underpinning assumptions, world views, and shared meanings which provide the substrate for large domains of human knowledge.

One example of knowledge-in-depth which comes to mind and which is currently a focus of policy activity relates to early childhood. The history of our understanding of human development during early childhood is a record of changing world

views — new views emerging and different views struggling for dominance. The legacies of these changes in world view do not disappear — they remain implicit in the research literature and continue to inform the ideas which emerge from it. For instance, studies of cognitive development during childhood, that is, studies of how thinking and knowledge develop during childhood, are now dominated by two approaches, the Piagetian and the Vygotskian, but resonances remain of other views, such as, behaviourism and nativism.

The behaviourist view is that children are blank slates which are *passively* moulded, or inscribed, by their experience — in this view, learning history is everything. Policy initiatives such as boot camps for juvenile offenders are based in part on a behaviourist perspective. The nativist position is that development of the individual child unfolds inexorably, guided by a biological programme, and that the child is *passively* changed in this process. Explanations of human behaviour which are based in an argument about genetic markers are essentially nativist. These views remain in play because they continue to provide some explanatory value and because there is evidence which supports them.

In contrast to the behaviourist and nativist positions, both Piaget and Vygotsky held the view that children are *active* in their development — that they are active participants in seeking out and constructing knowledge. In spite of this commonality, Piaget and Vygotsky held very different ideas about how cognitive development happens. Piaget advocated the idea that cognitive development in the individual child is a product of the *individual* mind; he saw children as essentially engaged in experimenting and making observations of their world. Children, according to the Piagetian approach, engage in making theories about how the world works. Vygotsky, on the other hand, viewed children's cognitive development as a deeply *social* process. According to Vygotsky, children's cognitive development is achieved through their interaction with other more knowledgeable members of their society. These different views have consequences for how one envisages early childhood education — including the current debate about language learning and baby DVDs!

Policy makers are experts in policy making and have knowledge-in-depth about policy, politics and government. Is it reasonable to ask that policy makers working in the early childhood domain also have knowledge-in-depth of these underpinnings of our knowledge of developmental psychology? In general, is it reasonable to ask that policy makers have knowledge-in-depth of the domains in which they make policy, or should they rely upon experts? If they should rely upon experts, then how does a policy maker — who is outside the community of experts in an area — evaluate the quality of evidence, or situate it within its larger context, that is, within the context provided by knowledge-in-depth? How does someone outside the community of experts even decide what counts as evidence? (This is often far from uncontroversial, especially in the social sciences and related domains.)

The contribution of evaluators — risks and opportunities

In 2005 Erik Arnold made some comments about the state of the practice of evaluation in Europe. Arnold wrote that:

The professionalism [of evaluators] is also a problem, because it involves the idea that evaluation is atheoretical and can therefore be done without reference to theory about the domains in which it works. As a professional group, evaluators need to make this claim, so that they can move among domains. And the claim is not absurd. A great deal can be done using a combination of deductive logic and a set of social-scientific and statistical tools. But it follows, first, that professional evaluators often do not make use of helpful theory in the domains where they work, usually because they neither understand nor are aware of it, and, second, they are unaware of the theoretical assumptions they build into evaluations. (Arnold, 2005, p. 35)

Looking around at evaluation in Australia, I think that these comments have some validity here, too. Many evaluation reports are built on a generic approach to thinking about evaluation — evaluation is presented as a theory-free zone — and, as Arnold

notes, this approach has great strengths. But it also has risks. Arnold mentions two: Our evaluations may be weaker because we do not use theory current in the domain in which we are working; and we ourselves may be unaware of the implicit theoretical assumptions we build into our evaluations. But, from the earlier discussion in this paper, it is clear that there is a third risk which should be added to these two. To the extent that Arnold's comments about Europe are correct for Australia, then it may be that neither evaluators nor policy makers necessarily possess and/or communicate knowledge-in-depth of the domain in which they are working. If this is true, then there is a problem for the interface between evidence-making and policy-making, since evaluators are often the evidence-makers who are in closest proximity to policy-makers. All of that said, it is also true that, in practice, many evaluators in Australia work within a specific domain and do not try to be completely generalist in applying their skills to any and all domains. This suggests that many of us probably have at least some knowledge of the domain in which we conduct evaluations, even if we currently do not communicate much of this in our reports.

Recently the OECD's Centre for Educational Research and Innovation (better known as CERI) produced a report with the title *Evidence in Education: Linking Research and Policy*; this report focused on issues surrounding the use of evidence in educational policy. In their chapter, Burns and Schuller wrote:

Ideas for education reform that originate from policy makers will be as varied as the system itself. Whether or not those ideas are based on evidence and research will depend, in large part, on the priority given to evidence-based policy research in education in that particular country.... *It will also depend on the ability of policy makers to hear and absorb the information that is being presented to them for use in evidence-based policy, and in their ability to understand the fundamentals of evaluation.* (Burns and Schuller, 2008, p. 21, emphasis added)

Now that's a very interesting sentence; in it we begin to see the kind of opportunities which we might forge for ourselves and our clients — win-win situations bringing advantages to all.

Developing our individual knowledge-in-depth about the domains in which we conduct evaluations, and communicating that knowledge-in-depth in a targeted fashion which is suited to the needs of our policy-maker clients, has several potential gains:

1. Our evaluations will be stronger because they are better fitted to the domain in which they take place.
2. We will be more able to analyse our own work, and so improve it and/or defend it.
3. Our policy-maker clients will receive reports which are richer and more likely to facilitate their continued making and refining of policy.
4. Evaluation reports might be less left-on-the-shelf and become more influential in the mix of needs and information which drive policy.

It is fair to say that evaluation is still an emerging profession. The time has come for us to begin tailoring our practice of evaluation to specific domains and establishing our specialist expertise within those domains. The time has come for us as professionals to build greater value into our evaluation reports. This will be good for our policy-maker clients, good for us, and good for the profession.

REFERENCES

- Arnold, E. (2005). What the evaluation record tells us about Framework Programme performance. Technopolis Group.
- Burns, T. and Schuller, T. (2007). The evidence agenda. In T. Burns and T. Schuller (eds.) *Evidence in education: Linking research and policy*, pp 15-32. OECD Publishing.
- House of Commons Science and Technology Committee. (2006). *Scientific advice, risk and evidence-based policy making, Document HC 900-1*. The Stationary Office Limited, London.
- OECD/CERI (2007). *Evidence in education: Linking research and policy*. OECD Publishing.
- Zimmerman, F., Christakis, D., and Meltzoff, A. (2007). Associations between media viewing and language development in children under age 2 years. *Journal of Pediatrics*, 151(4) 364-368.