Evidence-based policy: principles and requirements

Brian Head
University of Queensland

Abstract

Evidence-based policy (EBP) is an aspiration rather than an accomplished outcome. The advocates of EBP urge the incorporation of rigorous research evidence into public policy debates and internal public sector processes for policy evaluation and program improvement. The primary goal is to improve the reliability of advice concerning the efficiency and effectiveness of policy settings and possible alternatives. This is attractive to pragmatic decision makers, who want to know what works under what conditions, and also to those professionals concerned with improving information bases and improving the techniques for analysis and evaluation. Some concerns are raised by professionals whose knowledge-discipline or whose policy focus is not well served by quantitative analytical techniques, and who worry that important qualitative evidence may be overlooked. Scientific experts may reasonably disagree about methods, instruments and impacts. Whatever methodologies are employed, EBP requires good data, analytical skills and political support. Hence there are inherent limitations, even where government officials are able to draw on the results of reliable information and sound analytical skills. The politics of decision making inherently involves a mixing of science, value preferences, and practical judgments about feasibility and legitimacy. Outside the scientific community, the realm of knowledge and evidence is even more diverse and contested. Competing sets of evidence and testimony inform and influence policy. The professional crafts of policy and program development require ‘weaving’ these strands of information and values. The cutting-edge issues in modern EBP debates focus on problem-framing, methods for gathering and assessing reliable evidence, communicating and transferring knowledge into decision making, and evaluating the effectiveness of implementation and program delivery in complex policy areas.
2.1 Introduction

There are three crucial enabling factors that underpin modern conceptions of evidence-based policy (EBP): high-quality information bases on relevant topic areas, cohorts of professionals with skills in data analysis and policy evaluation, and political incentives for utilising evidence-based analysis and advice in governmental decision-making processes. The precursors of modern EBP thus have a long, if patchy, history over many decades, inspired by a desire to improve social, economic and environmental outcomes through the application of reliable knowledge. The story of EBP is as much about institutional development as about data and skills.

Australia, along with other prosperous Western countries, has developed a strong institutional foundation for nurturing EBP capacities. However, those enabling factors have developed unevenly, being more prominent in some periods than in others. For example, postwar reconstruction — a major theme in the federal government’s policy and planning concerns from about 1943 — arguably galvanised those factors. However, there was a political retreat from comprehensive ‘planning’ discourses under the Menzies government in the 1950s, although the infrastructure for data collection and skills development continued in various ways. Policy development and innovation, especially in social and urban policy, again became a strong theme in the early 1970s under a reformist federal government, and there was another boost from the mid-1980s with a stronger policy emphasis on economic productivity, regulatory liberalisation, and new approaches to social equity and environmental protection. As higher education expanded, the general culture of business and government became more favourable to the creation of ‘policy intellectuals’ in various institutional niches (Head 1988; Withers 1981).

Over time, specialised governmental organisations devoted to systematic data collection, the analysis of information and the evaluation of policy options, including the Productivity Commission, have grown in size and capability. In particular, long-term public investment in economic and social statistics, along with the development of more specialised units for policy and regulatory analysis, have provided a solid foundation for contemporary EBP capacities (Banks 2009). This investment has been driven by broad and diverse policy needs — for example, the need to understand and influence population trends, to assess and improve environmental sustainability, to plan and fund effective human services and social security, to provide better communications and transport infrastructure, to assess tax revenue capacities, and to meet the economic productivity challenges of international competition. (I omit here the highly distinctive research/policy needs of foreign policy, defence and intelligence organisations.) The Australian Government’s commitment to good data and sound analysis has also been reinforced by its growing involvement in international organisations (such as the
Organisation for Economic Co-operation and Development) and its endorsement of international agreements that require sophisticated reporting on comparative performance trends.

Beyond the sphere of the federal government, there have also been some important investments in EBP capacities within State governments, but on a much smaller scale and with generally lower levels of political support. Much of the impetus for States to invest in EBP capacity has been linked to their involvement with the powerful intergovernmental policy reform processes driven through the Council of Australian Governments and to a lesser extent other ministerial councils. Other new inputs to policy development have been associated with the recent proliferation of policy-oriented consultancy firms, and the emergence of independent think tanks (Marsh and Stone 2004) outside the public sector, sharpening the ongoing debates about the quality and timeliness of advice.

The United States has long been the major global location for policy analysis and evaluation professionals, both within government and in other policy-relevant sectors (see, for example, Lerner and Lasswell 1951; Nathan 1988; Wilson 1981). The refinement of methodologies for evidence-based assessment of program implementation, and for analysing alternative policy options, has been driven substantially by large cohorts of US scholars and policy managers. The mandating of particular forms of program appraisal as a condition of program funding has also proceeded further in the United States than in most other nations (Boruch and Rui 2008, Haskins 2006), although the practical experiences of evaluation remain diverse and somewhat fragmented across agencies and levels of government.

The British Government under Prime Minister Blair attempted to develop a coherent approach to policy development, championing EBP as a major aspect of the increased policy capability and the fresh thinking required by a reformist government (UK Cabinet Office 1999a, 1999b). This increased respect for research and evaluation was generally welcomed by policy researchers (for example, Davies et al. 2000), although some were troubled by the implicit preference for quantitative precision and for technical expertise over other forms of knowledge (for example, Parsons 2002, 2004). One of the positive outcomes has been a more comprehensive investment in policy-relevant research and a stronger commitment to evaluation.

In Australia, Prime Minister Rudd announced on 30 April 2008 that his government, not unlike the Blair government a decade earlier, saw a strong link between EBP and good governance:

A third element of the Government’s agenda for the public service is to ensure a robust, evidence-based policy making process. Policy design and policy evaluation should be driven by analysis of all the available options, and not by ideology. When preparing
policy advice for the Government, I expect departments to review relevant developments among State and Territory governments and comparable nations overseas. The Government will not adopt overseas models uncritically. We’re interested in facts, not fads. But whether it’s aged care, vocational education or disability services, Australian policy development should be informed by the best of overseas experience and analysis. In fostering a culture of policy innovation, we should trial new approaches and policy options through small-scale pilot studies.

Policy innovation and evidence-based policy making is at the heart of being a reformist government. (Rudd 2008)

These sentiments have been well received. However, the practical implications remain open to interpretation and debate, especially as to whether EBP in Canberra will entail an incremental approach building on best practice in professional analysis and advice, or will require greatly enhanced professional practices and wider adoption of specific skills (for example, cost–benefit analysis: Argyrous 2009). The initial lack of explicit guidance concerning preferred methodologies may have been a matter of either serious concern or great relief for different sections among the policy professionals. The overall level of commitment to investments in policy-relevant research, program evaluation and policy skills training in Australia has been disappointing, especially at State level. It remains to be seen whether the reinvigorated commitment to EBP will lead to measurably greater investment in policy research and evaluation over the coming years. In the following sections, I raise some basic issues about the political and institutional context in which EBP is pursued, as well as the internal debates about methods and reliable evidence.

2.2 Knowledge and rigour

The advocates of EBP urge the incorporation of rigorous research evidence into public policy debates and internal public sector processes for policy evaluation and program improvement. The primary goal is to improve the reliability of advice concerning the efficiency and effectiveness of policy settings and possible alternatives. The quest for rigorous and reliable knowledge, and the desire to increase the utilisation of rigorous knowledge within the policy process, are core features of the EBP approach.

Two main kinds of critical commentary have been expressed by observers and participants. The first can be termed ‘internal’ critical commentary, focusing on the suitability of various preferred methodologies for collecting, interpreting and applying evidence as a basis for understanding — and perhaps improving — particular programs or policies. The second can be termed an ‘external’ or contextual commentary, focusing on how and where the EBP contributions (based on rigorous evidence) can be most influential, and how they fit into the wider
picture of policy debate and evaluation — a public canvas largely painted by partisan viewpoints.

The quest for rigour is vital. There are many sophisticated sources of guidance for methodological questions of data validity, reliability and objectivity. Those who focus on these fundamental issues are usually specialists in the design of information capture and analysis (such as the Australian Bureau of Statistics and the Australian Institute of Health and Welfare) and/or specialists in the design of applied research on specific problems or programs. Two of the most widely debated matters in recent years are the significance of ‘qualitative’ evidence and a possible ‘hierarchy’ of reliability in different models of applied research. These are matters discussed by other papers and are noted here only briefly.

Concerns about the value of qualitative evidence stem from different research traditions in the social sciences. Some disciplines (for example, social anthropology and history) have usually tended to be concerned with accounting for the ‘experience’ of participants — meanings, motives, contexts — rather than seeking behavioural generalisations (which are more typical of quantitative approaches relying on economic and social statistics). Bridges are being built between the advocates on both sides. Program evaluation professionals tend to use mixed methods as appropriate. Large-$N$ qualitative studies are increasingly seen as open to the techniques of quantitative analysis. Longitudinal panel studies are an increasingly rich source of several types of evidence. In the United Kingdom, the early champions of very rigorous EBP quickly found it necessary (for example, Davies 2004) to soften the apparent bias towards randomised controlled trials (RCTs); and the central agencies have recognised that qualitative studies are important, provided they are conducted with appropriate methodological rigour (UK Cabinet Office 2003, 2008; UK Treasury 2007). Mixed methods are increasingly championed by analysts who are attempting to explain complex problems and assess complex interventions (for example, Woolcock 2009).

Turning to the related debate on ‘evidence hierarchy’, the underlying question is trust in the reliability of research findings. Some argue that there is a research-quality hierarchy, based on the types of methodological rigour used to design and interpret field studies. In particular, it is claimed that the RCT approach pioneered in medical research can and should be applied in the social sciences (Leigh 2009). A variant of this is the argument that single-study findings are misleading, and that a better understanding of causes and consequences emerges from ‘systematic reviews’ of all available research (Petticrew 2007; Petticrew and Roberts 2005), taking into account the rigour of the methods followed. The counterarguments turn on the difficulty of implementing RCTs in sensitive areas of social policy; the difficulty of transplanting quasi-experimental results to large-scale programs...
(Deaton 2009); and the tendency to downplay the knowledge and experience of professionals with field experience (Pawson 2006; Schorr 2003). It is also highly likely that politicians, policy managers, scientists and service users may have very different perspectives on what kinds of evidence are most trustworthy (see, for example, Glasby and Beresford 2006).

2.3 Knowledge for policy: how many lenses?

The movement to improve the evidence base (or bases) available for policy analysis and for program improvement is of crucial importance and is widely supported in many quarters. Encouraging organisational cultures that support more systematic evaluation of initiatives and interventions is also crucial. But building capability can be expensive. Moreover, providing transparent evaluations of program initiatives can be politically risky. Governments do not relish being exposed to strong public criticism for poor program outcomes or for pilot schemes that produce weak results. The culture of evaluation is best understood as a culture of learning, and therefore needs to be embedded as bipartisan good practice.

The knowledge base for EBP is diverse. Systematic research (scientific knowledge) provides an important contribution to policy making, and is undertaken in external institutions as well as in the public service. But science is only one of the inputs for EBP. The larger world of policy and program debate comprises several other types of knowledge and expertise that have legitimate voices in a democratic society. It has been argued elsewhere (Head 2008b; Shonkoff 2000) that these other ‘lenses’ or knowledge bases may include the following:

- The political strategies, tactics and agenda setting of political leaders and their organisations set the ‘big picture’ of priorities and approaches. The logic of political debate is often seen as inimical to the objective use of policy-relevant evidence: ideas and values are instead mobilised to support political objectives and to build coalitions of support; spin may become more important than accountability.

- The professional knowledge of service delivery practitioners and program coordinators is vital for advising on feasibility. They have crucial experience in service delivery roles and field experience in implementing and monitoring client services across social care, education, health care, etc. They wrestle with everyday problems of effectiveness and implementation, and develop practical understandings of what works (and under what conditions), and sometimes improvise to meet local challenges.
• In addition to the above institutional sources of expertise, the experiential knowledge of service users and stakeholders is vital for ‘client-focused’ service delivery. Ordinary citizens may have different perspectives from those of service providers and program designers; their views are increasingly seen as important in program evaluation for ensuring that services are appropriately responsive to clients’ needs and choices.

As illustrated in table 2.1, rigorous and systematic science seeks a voice in a competitive struggle for clarity and attention, jostled by many players in the wider context of public opinion and media commentary. To the extent that rigour is valued, it therefore needs to be protected by strong institutions and robust professional practices.

<table>
<thead>
<tr>
<th>Table 2.1 Types of knowledge relevant to evidence-based policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Political knowledge</td>
</tr>
<tr>
<td>Scientific rigorous knowledge</td>
</tr>
<tr>
<td>Professional–managerial knowledge</td>
</tr>
<tr>
<td>Client and stakeholder knowledge</td>
</tr>
</tbody>
</table>

The mass media and political culture

2.4 The policy process

The value proposition for EBP is that policy settings can be improved on the basis of high-quality evidence. How does reliable knowledge actually flow between producers and users? How strong are the channels and relationships that improve those flows? Unfortunately, the channels through which rigorous evidence may influence policy making are readily disrupted by external pressures, and therefore need specific care and attention (Landry et al. 2001; Lavis et al. 2003; Nutley et al. 2007). Supply-side provision of good research about ‘what works’ is not enough. Potential users will pay closer attention only when they are better aware of these inputs, understand the advantages and limits of the information, and are in a position to make use of the findings either directly or indirectly (Edwards 2004; Nutley et al. 2007). How can the social and institutional foundations for EBP be improved? What capabilities need to be built within and across organisations? Considerable research on such matters is already being undertaken across a range of social policy areas but cannot be summarised here in detail (for example, see Boaz et al. 2008, Dopson and Fitzgerald 2005; France and Homel 2007; Jones and Seelig 2005; Mosteller and Boruch 2002; Lin and Gibson 2003; and Saunders and Walter 2005).
A related matter is to determine at which points in the policy development and policy review ‘cycle’ EBP contributions (based on rigorous evidence) can be most influential. Based on the few available studies of policy development in Australia, it would appear that there are no general answers (Edwards 2001). The formal expectation might be that policy-relevant research about the effectiveness of various options (what works under what conditions) might be most closely linked into the evaluation phase of the policy cycle (for example, Roberts 2005). However, the notion of a rational and cyclical process of policy development, implementation and review does not correspond closely with political realities (Colebatch 2006). A more realistic and complex model is conveyed in a diagram published by the Scottish Executive (see figure 2.1), which allows for reiteration of process steps and continual processes of further consultation and gathering of evidence.

**Figure 2.1  The policy process**

Rigorous evidence can therefore be relevant at several points in the development and review processes. But not all matters are genuinely open to rethinking. Some areas of policy are tightly constrained by government priorities, electoral promises and ideological preferences. There is perhaps less scope for changing these as a result of evidence about ‘what works’. It is also useful to identify program areas that appear to be more settled than others over a period of time (Mulgan 2005). It is possible that evidence-based arguments about ‘fine-tuning’, based on careful research about effectiveness, might be more likely to gain traction in those areas if
they are away from the political heat. On matters of deep controversy, however, research findings are more likely to be mobilised as arrows in the battle of ideas, and sometimes in ways that the original authors may find distasteful. In this sense, the policy process is a patchwork quilt of arguments and persuasion (Majone 1989). However, policy adjustments, and opportunities for new thinking, can emerge in unexpected ways in response to incidents, crises and conflicts.

2.5 The tangles of complexity

Problems can be conceptualised at a variety of scales and with varying degrees of complexity. This is inherent in the nature of politics and policy debate. The scale or unit of analysis (for example, micro or macro level) and complexity (one issue or a nest of related issues) makes a big difference to how policy problems are framed, debated and researched. One feature of modern policy making is that a lot of ‘big’ problems are being addressed at the same time. Examples in Australia include specific programs to tackle Indigenous disadvantage, navigate the global financial crisis, respond effectively to the challenges of climate change, reform the health and education systems, and promote frameworks for social inclusion and nurturing early years development. Such large focus areas of policy attention pose challenges for evidence, analysis and recommendation. The political requirement for solutions will sometimes encourage broad-brush responses, rather than detailed bottom-up research as a basis for program trials and for careful evaluation prior to larger-scale implementation.

Some of these large problems have been termed ‘wicked’, owing to their resistance to clear and agreed solutions. These systemic and complex problems are marked by value divergence, knowledge gaps and uncertainties, and complex relationships to other problems (see figure 2.2, and discussion in Head 2008c and APSC 2007). It is not clear that traditional bureaucratic structures, or even the sophisticated managerial approaches of modern outcome-focused governments, are able to tackle these intractable problems successfully through standardised approaches.

One of the features of complex social problems is that there are underlying clashes of values, which are sometimes not adequately recognised and addressed (Schon and Rein 1994). Policy analysts have therefore tended to drift instinctively into two camps — one group tends to look for simple technical solutions (for example, inserting specific conditions into the funding for individual recipients of a program), whereas others seek to focus on identifying the underlying value conflicts as a basis for dialogue, mediation and conflict reduction prior to discussion of next steps toward solutions (Lewicki et al. 2003).
The attractiveness of recent behavioural approaches based on incentive theory (for example, Thaler and Sunstein 2008) is that judicious ‘nudging’ of citizens through incentives and penalties can potentially produce positive outcomes, with less need for intensive long-term case management or other expensive oversight and compliance mechanisms. In effect, the citizens are ‘nudged’ towards voluntary behavioural change arising from the ‘choice architecture’ embedded in the program design. Similarly, the attraction of quasi-market mechanisms for allocating scarce resources (such as irrigation water) and the perceived advantage of voluntary codes of conduct for industry is that detailed prescriptive regulation can be minimised (APSC 2009b).

The alternative widely favoured approach for addressing complex social problems is participatory collaboration, partnering and devolution (see APSC 2007, 2009a). The difficulties of such approaches are well known in terms of time, energy and ambiguity. Multiple stakeholders certainly complicate the challenges both of designing clear programs with defined roles and responsibilities, and of assessing the effectiveness of outcomes to be achieved through collaboration (Head 2008a). The social science challenges of evaluating complex programs are significant. Nevertheless, there is a very strong case for persisting in the face of complexity, since the underlying problems are of enormous importance for governments and citizens alike.
Some elements of these policy puzzles may be amenable to close scrutiny via rigorous appraisal and even through commissioning more RCTs. This is desirable. But the place of high-quality case studies in the broader context of complex policy challenges will need to be carefully contextualised. The professional crafts of policy and program development will continue to require ‘weaving’ together the implications of case studies with the big picture, and to reconcile the strands of scientific information with the underlying value-driven approaches of the political system.

References


Campbell Collaboration, http://www.campbellcollaboration.org/


Mulgan, G. 2005, ‘The academic and the policy-maker’, presentation to Public Policy Unit, Oxford University, 18 November.


Petticrew, M. 2007, ‘Making high quality research relevant and accessible to policy makers and social care practitioners’, presentation to Campbell Collaboration Colloquium, 16 May.


