

# **Is Evidence-Based Government Possible?**

**Jerry Lee Lecture 2004**

To be presented at the 4<sup>th</sup> Annual  
Campbell Collaboration Colloquium  
Washington D.C.  
19 February 2004

by

Philip Davies PhD  
Government Chief Social Researcher's Office  
Prime Minister's Strategy Unit  
Cabinet Office  
London  
SW1A 2WH  
England

[phil.davies@cabinet-office.x.gsi.gov.uk](mailto:phil.davies@cabinet-office.x.gsi.gov.uk)

**Foreword**

Evidence-based policy has become a major part of many governments' approaches to policy making and the machinery of government. The growing interest in, and practice of, evidence-based government in a number of countries has provided the inspiration for the title of this lecture. There are undoubtedly many competing ideas about evidence-based government and evidence-based policy, and many models from which we can learn a great deal. In this lecture I am mainly focussing on evidence-based government and evidence-based policy from a United Kingdom perspective for no other reason than it is the area in which I work and with which I am most familiar. This in no way implies that the UK's approach to evidence-based government is the only approach to this concept, or that my lecture fully captures what constitutes evidence-based government in the U.K. It is of necessity a personal view of how evidence-based government has been developed in the U.K. and of some of its constituent features. It goes without saying that this lecture does not represent an official view of the U.K. government.

**Abstract**

This lecture addresses whether evidence-based policy and evidence-based government is possible, and whether it is more than a rhetorical device. It attempts to define evidence-based policy and considers factors other than evidence that influence policy making and policy implementation. It also considers the types of evidence used by governments and the types of research that can contribute to that evidence. Some of the mechanisms that need to be in place for evidence-based government to occur are also discussed. The lecture concludes that evidence-based government is possible and is well established in the U.K. It argues that a broader conception of evidence is used by most government than by some academics, and that a wide range of methods for gathering and appraising evidence for government is required. Some implications for the Campbell Collaboration and the academic community are suggested.

*‘There is nothing a government hates more than to be well-informed; for it makes the process of arriving at decisions much more complicated and difficult.’*

John Maynard Keynes

## **Introduction**

Reforming and modernising the machinery of government has been a central feature of the Blair administrations of 1997 and 2001. Part of this modernisation and reform has been a commitment to evidence-based policy. The *Modernising Government White Paper* (Cabinet Office 1999a), for instance, stated that government policy must be evidence-based, properly evaluated and based on best practice. A report from the Cabinet Office Strategic Policy Making Team on *Professional Policy Making for the Twenty-First Century* (Cabinet Office, 1999b) suggested that “policy making must be soundly based on evidence of what works” and that government departments must improve their capacity to make use of evidence. A later Cabinet Office report, *Better Policy Making* (Cabinet Office 2001a), surveyed policy makers across government and concluded that policy making was “more informed by evidence” than had previously been the case. It cited the reviewing of existing policies, the commissioning of new research, the piloting of new initiatives, and the evaluation of new policies as examples of the greater use of evidence in policy making across the U.K. Government.

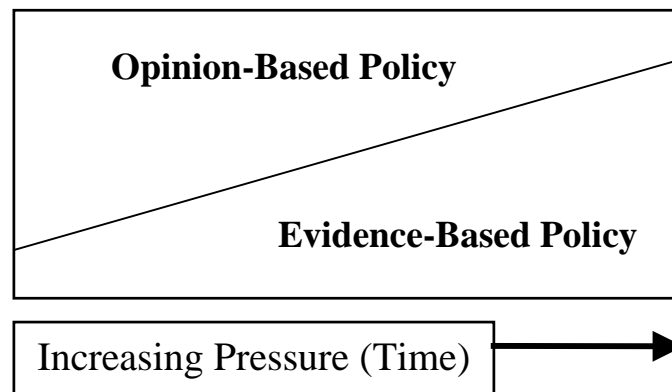
This paper attempts to define evidence-based policy more precisely and will consider the role of evidence in the policy making process. It argues that evidence-based government is possible and provides some examples of this from a United Kingdom perspective. However, the paper also suggests that there are things other than evidence that contribute to policy making, and that there are competing types of evidence that are used by policy makers and those responsible for policy implementation and delivery. The challenges that this raises for government research, evaluation and analysis are also considered. Finally, this paper reviews what is known about the mechanisms that need to be in place for evidence-based policy to be effective.

### What is Evidence-Based Policy?

Evidence-based policy has been defined as an approach that “helps people make well informed decisions about policies, programmes and projects by putting the best available evidence from research at the heart of policy development and implementation” (Davies, 1999a). This approach stands in contrast to opinion-based policy, which relies heavily on either the selective use of evidence (e.g. on single studies irrespective of quality) or on the untested views of individuals or groups, often inspired by ideological standpoints, prejudices, or speculative conjecture. Gray (1997) has suggested that there is a new dynamic to decision making in health care and other areas of public policy whereby the speculation of opinion based policy is being replaced by a more rigorous approach that gathers, critically appraises and uses high quality research evidence to inform policy making and professional practice. A graphical representation of this new dynamic is presented in Figure 1.

**Figure 1**

#### **The Dynamics of Evidence-Based Policy**



(Adapted from Gray, 1997)

### Methodological Problems and Social Scientific Knowledge

Proponents of evidence-based policy and practice acknowledge that not all research is of a sufficient quality to form the basis of sound policy making (Davies, Nutley and Smith, 2000). Many research studies are flawed by unclear objectives, poor research designs, methodological weaknesses, inadequate statistical reporting and analysis, selective use of data, and conclusions that are not supported by the data

provided. In light of these shortcomings evidence-based policy requires a more systematic approach to searching for appropriate evidence, the critical appraisal of studies that are identified, and a balanced understanding of what the research evidence is saying and of its strengths and weaknesses (Davies, 2003). The potential contribution of the Campbell Collaboration to this aspect of evidence-based policy is clearly considerable.

A further problem for evidence-based policy is the *uncertainty* of social scientific knowledge, and the different status of different fields of knowledge. Mulgan (2003) has suggested that the latter runs on a continuum from fields of knowledge that are well established and almost like a 'normal' science to those where knowledge is inherently novel, such as global governance, regulation of biotech, and e-government. These problems mean that in many areas of policy making there is either little or no valid social scientific evidence, the consequence of which is a knowledge vacuum that other types of evidence can fill until new sound research evidence can be established.

### **Factors Other Than Evidence**

The above definition of evidence-based policy and practice has been challenged on the grounds that policy making involves factors other than evidence, and that to give such a central place to research evidence is misplaced. This objection has some merits and warrants further consideration.

#### *Experience, Expertise and Judgement*

One group of factors that clearly influence policy and practice is the experience, expertise and judgement of decision makers. These factors often constitute valuable human and intellectual capital and include the tacit knowledge that has been identified as an important element of policy making (Nutley, Walter and Davies, 2003). Such influences may, or may not be informed by sound evidence. Indeed, judgement based on experience and expertise may be of critical significance in those situations where the existing evidence is equivocal, imperfect, or non-existent

(Grimshaw, *et al*, 2003). Consequently, a major goal of evidence-based policy is to ensure that policy making *integrates* the experience, expertise and judgement of decision makers with the best available external evidence from systematic research.

### *Resources*

Policy making and policy implementation take place within the context of finite (and sometimes declining) resources. This means that policy making is not just a matter of ‘what works’, but what works at what cost and with what outcomes (both positive and negative). This requires sound evidence not only of the cost of policies, programmes or projects, but also the cost-effectiveness, cost-benefit, and cost-utility of different courses of action. Economists have well developed methods for appraising, analysing and evaluating such factors and there are a number of sources of government guidance on how to do this (HM Treasury, 2003; Cabinet Office, 2001b, 2003a; ODPM 2000).

### *Values*

Policy making also takes place within the context of values, including ideology and political beliefs. Political ideology is a major driving force of policy making and is in no way made redundant by a commitment to evidence-based policy. Political ideologies have always been subjected to close critical appraisal and analysis using both philosophical and empirical methods. The tension between values, ideology and beliefs on the one hand, and sound empirical evidence on the other, is the very stuff of contemporary politics in open democratic societies, and is unlikely to disappear because of the advent of evidence-based policy. Indeed, evidence-based policy can itself be seen as a political ideology, representing the case for empirical demonstration alongside more theoretical approaches to political discourse and action.

### *Habit and Tradition*

Habit and tradition are also important influences on policy making. Political institutions such as Parliament, the Civil Service and the Judiciary are steeped in

traditional and habitual ways of doing things, many of which defy rational explanation in the twenty-first century. This situation is not confined to the United Kingdom and is often reinforced by rituals and ceremonial procedures that are deeply engrained in the fabric of political life. Changing traditional and habitual ways of doing things to accommodate the forces of rationality and modernity presents a major challenge for evidence-based policy and practice. Some adaptation in the opposite direction may also be necessary.

#### *Lobbyists, Pressure Groups and Consultants*

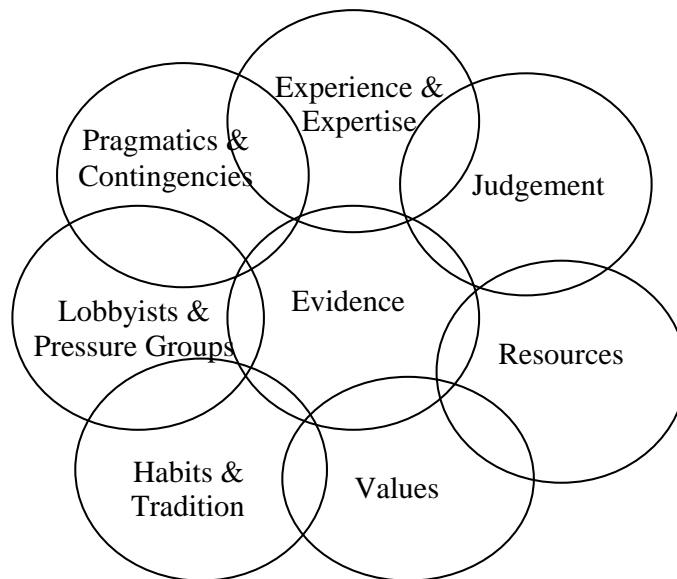
The lobby system, pressure groups and consultants are other factors competing with evidence to influence policy making and policy implementation. Think-tanks, opinion leaders and the media are other major influences. The ways in which these groups work to influence policy can be under-estimated and misunderstood by proponents of evidence-based policy and practice. It is not that these groups fail to use evidence to promote particular policies, programmes or projects. Rather, it is that such evidence is often less systematic, and more selective, than that used by supporters of evidence-based policy and practice.

#### *Pragmatics and Contingencies*

Other factors that influence policy making and policy implementation are the sheer pragmatics of political life such as parliamentary terms and timetables, the procedures of the policy making process, the capacities of institutions, and unanticipated contingencies that arise such as the Fuel Crisis of 2000 and the Foot and Mouth outbreak of 2001. These factors need not be the enemy of evidence-based policy and practice. First, evidence-based policy is a strategic as well as an operational activity, and part of its role is to build an evidence-base for future generations of policy makers and practitioners. Second, evidence-based policy and practice should be the first line of response to unanticipated events in the sense of identifying what is already known about the problem and what is not. This has been recognised by the public inquiries into both the Foot and Mouth outbreak (Royal Society, 2002) and the BSE crisis (Phillips Inquiry 2001).

A graphical representation of the features that influence policy making in government is presented in Figure 2. This is a provisional model that is subject to further development and specification.

**Figure 2**  
**Factors Influencing Policy Making in Government**



### **There Are Different Types of Evidence**

#### *Systematic Reviews*

The evidence-based policy movement, and the Campbell Collaboration especially, has built its claim to influencing policy and practice on the basis of using research evidence that has been *systematically* searched, critically appraised, and rigorously analysed according to explicit and transparent criteria. Systematic reviews and meta-analyses of existing evidence are accorded such a high premium amongst proponents of evidence-based policy and practice because they overcome the shortcomings of single studies (Cooper and Hedges, 1994; Davies, 2003). Single studies can provide an unbalanced and unrepresentative view of the total available evidence on a topic or policy issue. This is because they are almost always sample-specific, time-specific, and context-specific. Also, some single studies lack methodological rigour or are not undertaken to high quality standards. Such studies should not be included in the evidence base for policy making or practice.



Systematic reviews, by contrast, use explicit and transparent quality criteria, and rigorous standards for searching and critical appraisal, in order to establish ‘the consistencies and account for the variability of similar appearing studies’ (Cooper and Hedges, 1994:4). By only accumulating evidence that can warrantably be accumulated, and identifying studies that are sample, time, or context specific, systematic reviews are able provide generalisations, and the limits of these generalisations, amongst existing research evidence.

Rapid evidence assessments and interim evidence assessments are being developed in the UK government to provide real time research synthesis as a way of helping policy makers use existing research evidence. These methods of harnessing existing evidence are used with caution and with the proviso that the fully developed systematic review may change the balance of available evidence and the implications for policy and practice.

Much of the evidence that is used in policy making, however, is either less systematically gathered and appraised than the evidence-based policy movement would propose, or is generated by expert opinion, or both. Policy makers and other users of evidence may need to be made aware of what systematic reviews offer, and what they should be asking of them. This is sometimes referred to as developing the ‘intelligent customer’ role in policy making and policy implementation. There is a parallel need for producers of systematic reviews to better understand the evidence needs of policy customers and to produce reviews and other types of evidence that meet these needs. This might be referred to as the ‘intelligent provider’ role.

### *Single Studies*

Single studies are more commonly used than systematic reviews to support government policy and practice. Indeed, the vast majority of the research undertaken by, or on behalf of, the UK government consists of single studies, often

without any accumulation of existing evidence using systematic review methodology. If undertaken to the highest possible standards, single studies can provide valuable and focussed evidence for particular policies, programmes and projects in specific contexts. Unlike systematic reviews, however, single studies are less able to say much about the variability of populations, contexts and conditions under which policies might work or not work.

### *Pilot Studies and Case Studies*

Pilot studies and case studies are other sources of evidence for policy making and policy implementation. A recent review of pilots by the UK Cabinet Office (2003b) identified impact pilots, process pilots and phased implementation projects, each using a combination of experimental, quasi-experimental and qualitative methods, as well as case studies, to help guide policy makers and policy making. The Cabinet Office report recommended that “the full-scale introduction of new policies and delivery mechanisms should, wherever possible, be preceded by closely monitored pilots” (Cabinet Office, 2003b:5).

It is sometimes argued that the tight timetables and schedules of the policy making process make it impossible for systematic reviews, single empirical studies, pilots or case studies to be undertaken before rolling out a policy, programme or project. This reasoning is often deployed to justify the use of whatever evidence is readily available, regardless of its scientific quality or source. Such urgency and rapidity of action may be understandable, especially in the absence of a well established evidence base for many areas of public policy, but it is short sighted and possibly counter productive. Evidence that is selective, and not subjected to careful critical appraisal and risk assessment, can often lead to inappropriate courses of action which cause more harm than they are intended to prevent. The Dangerous Dogs Act is a case in point as is the imprecise use of evidence in the case of the BSE outbreak (Phillips, 2000) and the Foot and Mouth crisis (Royal Society, 2002).

Where pilots are used to test policies it is important that they are completed and that

lesson are learned before more widespread implementation. The Cabinet Office Review of Pilots recommended that:

Once embarked upon, a pilot must be allowed to run its course. Notwithstanding the familiar pressures of government timetables, the full benefits of a policy pilot will not be realised if the policy is rolled out before the results of the pilot have been absorbed and acted upon. Early results may give a misleading picture.

(Cabinet Office, 2003b: 5; Recommendation 6)

### *Experts' Evidence*

Expert opinion is also commonly used to support government policy and practice, either in the form of expert advisory groups or special advisers. Using experts as a basis for policy making and practice, however, again raises the problems of selectivity of knowledge and expertise, as well as ensuring that the expertise being provided is up to date and well grounded in the most recent research evidence. In the case of the BSE outbreak, for instance, the Phillips Inquiry identified the incorrect understanding by experts of the nature, cause and transmissibility of the disease. This led to an imprecise estimation of the likelihood of BSE being transmitted to humans and the subsequent inappropriate reassurance of the public by Ministers (Phillips Inquiry, 2001). The Foot and Mouth crisis of 2001 was hindered by experts' competing views about the nature and spread of the disease, and about the best ways of dealing with it. The Royal Society Inquiry (2002) into the Foot and Mouth outbreak identified the need for a better evidence base using sound real time data based on field epidemiology, mathematical modelling, and valid diagnostic tests and techniques.

Huw Davies (2000) has characterised many experts' panels with the acronym GOBSATT – good old boys sitting around talking turkey. The UK Office of Science and Technology (2001) has published guidelines for the selection and running of UK Government Advisory Groups to ensure that no such characterisation can be made of experts' committees in UK government. Ensuring that government experts are up to date with existing evidence, and with the existing *uncertainty* of scientific knowledge, is one feature of these guidelines.

### *Internet Evidence*

The internet age has brought a revolution in the availability of information and knowledge. Most, though not all, government departments in the UK have desktop access to the internet and some departments, though not all, have access to social science and political science databases. It is anticipated that all government departments will have both internet and database access within the near future. This means that there is uneven access across UK government departments to these important sources of potential evidence

Not all of the information available *via* the internet, however, is of equal value or quality. Many sites provide ‘evidence’ that is either scientifically or politically biased, or both. The uncertain scientific and political basis of much of the information and knowledge on the internet makes it difficult to be assured that it meets the required quality to be counted as sound, valid and reliable evidence. This makes it all the more important for government analysts and the wider academic community, including the Campbell Collaboration, to ensure that such information is critically appraised and scientifically assured before it is used as evidence for policy making purposes.

### **Different Types of Research Evidence**

There are many types of research evidence that can and should be used for evidence-based policy and practice. Privileging any one type of research evidence, or research methodology, is generally inappropriate for evidence-based government. The guiding principle for the types of evidence that are appropriate for policy making and implementation is: ‘what is the question?’ (Greene, Benjamin and Goodyear, 2001).

### *Impact Evidence*

The Campbell Collaboration is leading the way in ‘preparing, maintaining, and disseminating systematic reviews of *the effectiveness of interventions* in education,

crime and justice, and social welfare' (Davies and Boruch, 2001). Such reviews are mainly (though not exclusively) concerned with the *impact* of policy on outcomes, and are generally best served by studies that use experimental and quasi-experimental research designs with good counterfactual measures. The lead that the Campbell Collaboration and others have taken on this front is commendable and is leading to a greater awareness, and a higher quality, of studies on the likely impact of interventions.

### *Implementation Evidence*

Governments, however, are not just interested in the effectiveness of the *outcomes* of interventions; they are equally interested in the effectiveness of *the implementation and delivery* of policies, programmes and projects. These two types of effectiveness are indeed closely linked to each other. The importance of effective implementation and delivery has been highlighted in the U.K. since the General Election of 2001, when the reform and delivery of public services became the defining theme of the second Blair administration.

Experimental and quasi-experimental research designs can greatly help implementation and delivery issues by bringing a degree of comparative rigour to different modes of practice. Effective implementation and delivery, however, also requires high quality qualitative data using in-depth interviews, focus groups, other consultative methods (such as the Delphi and Nominal Group methods), observational methods, participant-observation methods, and social surveys. The UK Cabinet Office's *Quality of Qualitative Evaluation* framework (Spencer *et al*, 2003) is one contribution to ensuring that qualitative research is undertaken to agreed high quality standards. This and other developments, such as work on meta-ethnography (Britton *et al*, 2002; Campbell *et al*, 2003) and on including qualitative data in systematic reviews (Dixon-Woods, 2001; Harden *et al*, 2003), will enhance the synthesis of evidence from qualitative studies.

A recent review of the evidence on effective implementation, however, has described the field as 'imperfect' and often inconclusive (Grimshaw *et al*, 2003).

There is a very strong need for more and better implementation studies that can identify the particular conditions under which successful implementation and delivery takes place, or fails to take place, as well as those conditions that are more generalisable.

#### *Descriptive Analytical Evidence*

Another important type of research evidence for government comes from descriptive surveys and administrative data about the nature, size and dynamics of a problem, a population, sub-groups, or social activities. Cross sectional, time-series and comparative data on a wide range of variables are regularly collected and used by Governments using sophisticated descriptive and analytical methods. Government work on Strategic Audit and on benchmarking countries' performance and social changes over time (e.g. Cabinet Office, 2004) use descriptive analytical evidence extensively. Such data are also used for process and outcome measures in experimental and quasi-experimental studies, as well as for comparative descriptive purposes.

#### *Public Attitudes and Understanding*

Research evidence on the attitudes, values and understanding of ordinary citizens is very important for effective government. This is much more than government by opinion poll and goes to the point that policies that are too far removed from the grain of public values simply will not work. In the U.K., the Poll Tax and the approach of the Child Support Agency are cases in point. By contrast, the Department of Social Security's (now Department of Work and Pensions) programme of research on 'Attitudes to Welfare Reform' (Williams *et al*, 1999) was highly instrumental in framing the design and delivery of welfare-to-work policies. Citizens' perceptions, experiences and understanding of policy are generally best addressed using qualitative research designs and social survey methods.

#### *Statistical Modelling*

Statistical modelling also plays a very important part in the evidence base for

government. Such modelling uses linear and logistic regression methods, and assumptions about policy scenarios that need to be manipulated. The more these assumptions, and the variables used in statistical modelling, are based on sound empirical evidence, the greater will be the precision and external validity of such analysis.

### *Economic Evidence*

Other types of evidence that are routinely used by governments concern the cost, cost-benefit and cost-effectiveness of policies. Such evidence uses economic appraisal and evaluation methods, including econometric analysis and modelling, and is a central part of most governments' evidence base. Economic appraisal and evaluation is increasingly linked to the implementation and delivery agenda of governments (such as the biennial Spending Reviews of the UK Government), so that cost-effective and cost-beneficial interventions are rewarded with government funding, and cost-ineffective or non-beneficial programmes are not.

Performance Management in government means that target setting and chasing are increasingly being used as a way of establishing whether governments' goals are being met. Much of the criticism of targets concerns their internal and, especially, external validity, and the top-down ways in which they are often determined and imposed upon front line staff. The need for more sensitive and appropriate targets in many areas of government is compelling, and this requires both summative and formative evidence using experimental and non-experimental research methods.

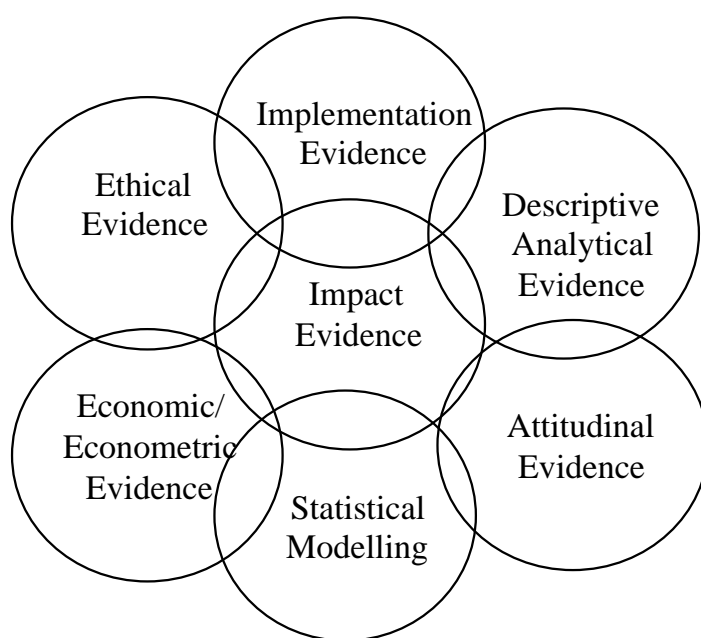
### *Ethical Evidence*

Governments also make daily decisions that involve trade-offs between one policy and another, or one group and another. This can mean withdrawing a programme or service from one group of people in order to provide a more cost-beneficial programme for another group. Such action requires evidence of relative effectiveness, relative costs, people's perceptions and experiences, and of the social

justice and ethics of doing so. Decisions about the latter often require evidence from social ethics, such as Rawls' (1972) *Theory of Justice*, and public consultation.

In summary, evidence-based policy and practice uses a range of types of research evidence, and is usually guided by the questions being posed rather than by any one type of research evidence. A graphical representation of the different types of research evidence used in government is presented in Figure 3.

**Figure 3**  
**Types of Research Evidence**



### **The Challenges for 'Intelligent Providers'**

#### *Multi-Method Capability*

The different types of research evidence that support evidence-based policy and practice require a government social research service, including that provided by external contractors and agencies, to have multi-method capability. A balance is required between social researchers having a *general* understanding and competence of the full range of research methods used in evidence-based policy making, and *specialist* skills and in-depth capability with particular methods. It is therefore important for social researchers who work to support evidence-based policy and



practice to have professional development opportunities in research methods and analysis.

### *Training and Professional Development*

The Government Chief Social Researcher's Office (GCSRO) in the UK, in collaboration with a number of external partners, has developed a programme of training and professional development for government social researchers that provides opportunities for updating research skills at the generalist and specialist levels (see [www.policyhub.gov.uk](http://www.policyhub.gov.uk) for details). It also provides guidance notes on social research methods for policy evaluation and analysis that are updated regularly (see the *Magenta Book* at [www.policyhub.gov.uk](http://www.policyhub.gov.uk)). The development of joint training for researchers and policy makers is also necessary if effective collaboration between 'intelligent providers' and 'intelligent customers' is to be achieved. There is clearly a need for greater provision of such training and professional development by groups such as the Campbell Collaboration.

### *The 'Challenge Function'*

The UK Government's review of the analytical needs of government to support evidence-based policy making (Cabinet Office, 2000) suggested that the central government departments (Cabinet Office and H.M. Treasury) should undertake a 'challenge role' with government departments. This would "review departmental analytical strategies to identify deficiencies, gaps and overlaps" in their analytical capability. An initial opportunity for such a challenge was provided by the Evidence-Based Policy Fund, which provided central resources to identify gaps and deficiencies in sound evidence for policy making. The subsequent Spending Review procedures, whereby government departments are centrally reviewed every two years to determine whether their spending targets have been met, is another means of engaging in this challenge function.

There is another sense in which government and external researchers can perform a 'challenge function' and that is by challenging the theoretical, analytical and

empirical basis of government policies, programmes and projects. The contribution of theories of change approaches to evidence-based policy making is clearly relevant here (Chen, 1990; Weiss, 2001; Rogers *et al*, 2000). Where social researchers have doubts about the evidence to support a particular policy or programme it is appropriate for them to challenge policy officials and, more constructively, to search and critically appraise the available research evidence on the topic in question. Hal Gehman, who headed the NASA investigation into the Columbia space shuttle disaster in 2003, has noted the need for a similar challenge function in that investigation. Gehman noted the ‘absolute certainty’ with which NASA scientists and engineers went about analysing the disaster. Gehman writes:

Now when I hear NASA telling me things like “Gotta be true!” or “We know this to be true”, all my alarm bells go off...Without hurting anybody’s feelings, or squashing people’s egos, we’re having to say, “We’re sorry, but we’re not accepting that answer”.

(Langewiesche, 2004:30)

The sooner a broad evidence-base for public policy is established, based on systematic reviews and high quality analysis, the easier it will be for such a challenge function to be undertaken effectively by social researchers, and for policies and programmes to be developed around sound evidence.

### *Improving Accessibility of Evidence*

A further challenge for social researchers is to make the findings of social research accessible to the policy making community. Too often social research is presented in an unclear way with as much, if not more, emphasis given to the caveats and qualifications of research findings (the ‘noise’ of social research) than to the message and implications of these findings for policy and practice (the ‘signal’). Government social researchers often need to ‘translate’ social science evidence into a language that is useful to the users of evidence, without distorting or misrepresenting the research evidence. It is quite acceptable for researchers and analysts to conclude that the available research evidence is inconclusive, or of an insufficient quality to provide a basis for policy making. It is not appropriate merely

to conclude that the problem is “very complex” (which is usually already recognised), or that “more research is needed” unless this is focussed and specific.

### **Mechanisms for Getting Evidence Into Practice**

There has been a considerable amount of work in recent years on how to get evidence into practice (Stocking 1992; Lomas, 1993; Davies, 1999b; Nutley and Davies, 1999; Davies, Nutley and Smith, 2000; Nutley, Walter and Bland, 2002; Nutley, Walter and Davies, 2003; Grimshaw *et al*, 2003; Davies, 2004,). This literature has identified the following factors:

#### *Integrating Research into Professional Competence*

A distinction can be made between people who are *users* of research and those who are *doers* of research. Whilst it may be unrealistic, and even undesirable, for professional decision makers and practitioners to be competent doers of research, it is both reasonable and necessary for such people to be able to understand and use research in their professional practice. Integrating research into practice is a central feature of professions such as medicine, law and engineering. Knowing about the different kinds of social, economic and policy research that are available, and how to gain access to them and critically appraise them, is an increasingly necessary skill for professional policy makers and practitioners to have. Without such knowledge and understanding it is difficult to see how a strong *demand* for research evidence can be established and, hence, how getting research into practice can be enhanced. Joint training and professional development opportunities for policy makers and analysts may be one way of taking this forward and for matching strong demand with a good supply of appropriate evidence.

#### *Ownership of the Evidence*

A closely related issue is getting policy makers and practitioners to own the evidence needed to support and implement policy effectively. This stands in contrast to a position where evidence is solely the property and domain of researchers or, perhaps even worse, managers and bureaucrats who try to impose

less than transparent evidence upon practitioners and front line staff. Ownership of the best available evidence can enhance its use to make well informed and substantiated decisions.

### *Getting Appropriate 'Buy-In'*

The literature cited above suggests that getting policy makers and practitioners to own and use evidence also involves getting commitment and buy-in at the most appropriate levels. In central government this usually means getting Ministers and senior policy officials to sign up to the ownership of a project *and* the evidence that goes to support it. This in turn means a commitment to use findings that are contrary to expectations, and not to continue with a policy, programme or project if the available research evidence indicates that this is ineffective. At the level of 'front line' service delivery it means getting key decision makers to 'own' and champion the evidence that supports good practice (Davies, 1999b, 2004). This is most likely to take place, and most likely to be effective, in organisational structures that are non-hierarchic, open and democratic (Dowd, 1994; Martin, 1997).

### *Shared Notions of Evidence*

Evidence is more likely to be used in the policy making process if there is agreement between policy makers and researchers, and within the research community, as to what constitutes evidence. The disputes between researchers about the superiority or inferiority of quantitative versus qualitative studies, or experimental versus experiential research designs, can lead to no useful evidence being produced, or to evidence that is technically very good but of little use to policy makers or anyone else. In the meantime, there are plenty of other sources of evidence – from lobbyists, pressure groups, consultants, the media etc. – that are less thorough but more readily available to policy makers. It is not surprising that such evidence is often more successful in finding its way into policy making.

The message is clear. Researchers must work with policy makers to more fully understand the sorts of questions they need to answer, and to agree on the most

appropriate evidence that will help solve policy problems. This means having a *strategic* approach to policy development and, wherever possible, *integrated* teams of policy officials, researchers, specialist consultants, and people who have to implement and deliver front line services. An example of such a team was provided at the design stage of the Employment Retention and Advancement (ERA) demonstration project (see below). Within government the term ‘bedded out’ is sometimes used to refer to the integration of researchers and other analysts into policy making teams. The more that external researchers can be ‘bedded out’, or ‘bedded in’ to policy making and service delivery teams the greater the likelihood of developing shared notions of evidence and of getting research into practice.

### *Incentives to Use Evidence*

The evidence on how to get research into practice repeatedly shows that practitioners need incentives to use evidence and to do things that have been shown to be effective. This also means not doing things that have been shown to be ineffective or even harmful. At the level of central government departments in the U.K. Public Service Agreements (PSAs) and Service Delivery Agreements (SDAs), coupled to the biennial Spending Reviews by HM Treasury, provide something of an incentive to establish evidence of effectiveness and efficiency. At the local level, the devolution of budgets to front line agencies and decision making bodies such as hospital trusts, primary care teams, local education authorities and school governors, has provided a similar incentive to summon and use sound evidence in resource allocation and service development. Targets set by central government, or even by more local agencies, however, may be too blunt and too insensitive an instrument to act as effective incentives to use appropriate evidence. There is always the possibility with target setting and target chasing of Type I errors (achieving an inappropriate target) and Type II errors (not achieving a target that is appropriate). Finding more sensitive and effective ways to incentivise people to use appropriate evidence is a major challenge for evidence-based government, and requires research evidence on how to change human behaviour. Such evidence is possibly the Achilles heel of social and behavioural science.

### *Availability of Sound Evidence*

Implicit in the whole question of how to get research evidence into practice is the existence and availability of sound evidence. The development of research synthesis by groups such as the Cochrane and Campbell Collaborations, the EPPI Centre, and the ESRC Evidence Network, has shown that there is often a lack of sound, conclusive evidence even when there has been considerable research activity on some topic or problem. Such conclusions need not be nihilistic; they can lead to the development of primary research that is well designed and able to fill the gaps in the evidence base. If the Campbell Collaboration needs any impetus to drive forward its mission it must be the urgent requirement of policy makers for high quality reviews of what we already know, and a clear indication of what new research is required to establish a sound evidence base for policy and practice.

### **Examples of Evidence-Based Government**

Some of the machinery of government already mentioned, such as the use of delivery and service agreements, national and local targets, and biennial spending reviews are ways in which evidence-based government has been developed in the U.K. For many researchers and academic observers, however, such approaches may not constitute evidence-based government so much as accountancy-based government. For this audience, evidence-based government usually means investing in research and analysis and using this as a basis for policy making.

The UK Government spends over £150 million a year on social and economic research, in addition to the £50 million it spends via the Economic and Social Research Council. Even greater sums of money are spent via the other research councils such as the Medical Research Council (MRC) and the Engineering and Physical Science Research Council (EPSRC), much of which provides evidence for policy making. There are approximately 4000 social researchers and other analysts working in government departments and agencies, 1000 of whom are government social researchers. With these sorts of money and numbers of people working on

government research and analysis, it would be a sorry state of affairs if the UK government was not evidence-based to some degree.

There are many examples in the UK of evidence-based government and evidence-based policy as defined in this paper, far too many to chronicle in one paper. Some notable examples that would meet the standards outlined in this paper include the Sure Start programme, the Educational Maintenance Allowance, the Connexions programme (DfES), many of the New Deal employment and welfare-to-work programmes (DWP), the New Deal for Communities, much of the Neighbourhood Renewal programme, the Home Buying and Selling policy (ODPM) and many policy pilots (Cabinet Office, 2003b). Other examples include the work undertaken by government departments to better understand the nature, size and distribution of problems that policy seeks to solve. Work on strategic audit and on benchmarking UK social life, social change and social institutions against those of other countries are other examples of evidence-based government in the United Kingdom. So too is the attitudinal work undertaken for and by government to establish public understanding and perceptions of policy.

#### *Employment and Retention Advancement (ERA) Demonstration Project*

One example of good evidence-based policy making is the design of the Employment and Retention Advancement (ERA) Demonstration project (Cabinet Office, 2003c). The ERA project design team was asked to identify the most effective means of retaining in the workforce low paid workers who had recently left welfare, and the most effective ways of advancing them in the labour market. A number of policy options were suggested at the outset and the team was asked to design an evaluation of the final policy using randomised controlled trial methodology. The project team was based in the Cabinet Office in order to provide an 'off-line' opportunity, and to work in a cross-cutting and cross-departmental way.

The design team consisted of policy officials, policy implementation staff,

researchers and specialist consultants from research organisations in the UK and the USA. This allowed a high degree of integration of policy making, policy implementation and policy research and evaluation. It also allowed time and resources to gather the best available evidence on effective interventions for the retention and advancement of people on the margins of the labour force.

An extensive, though not systematic, review of the evidence was possible due to the considerable work that had already been undertaken on employment and retention policy by research organisations such as the Manpower Development Corporation (MDRC), the Social Development Research Corporation (SDRC) and Mathematica. Additional qualitative evidence (in-depth interviews) and survey evidence was commissioned on Job Centre clients' views of the transition from welfare to work, and on employers' view of retention and advancement issues. Qualitative evidence was also summoned from representatives of Job Centre Plus offices, which was the agency that would deliver any new policies. This provided vital evidence about the local contexts in which the policy would operate, and about practical implementation issues of delivering new services within the context of a randomised controlled trial.

The review work and the qualitative and survey evidence narrowed the policy options down to the three that were most likely to be effective: an extended personal adviser service, tax incentives, and skills' development payments. Identifying appropriate and sensitive outcome measures, using both existing administrative data and primary survey data, required discussion with the project's advisory and steering groups, and careful negotiation with the relevant departments and agencies. Working with the steering group also provided the necessary 'buy-in' from Ministers, policy officials, and other government analysts at the appropriate levels.

A detailed and sensitive cost-benefit study is an essential part of the trial, the design of which was undertaken by a consultant economist and other experts in economic evaluation. It was politically important to test these interventions in three of the



constituent countries of the United Kingdom (England, Wales, and Scotland), and this enhanced the external validity of the demonstration. Liaison with front line staff (who would deliver the policy) in the six experimental and six control sites was another important part of the policy development.

The design of the ERA Demonstration project is an unusual, if not unique, approach to policy development in the UK context. It is unlikely to be used very often for policy development, implementation and evaluation if only because of the length of time involved – one year for the design phase and five years for full evaluation (with milestone data being produced along the way) – and the resources that were committed to it. Nonetheless, it does provide an example of how evidence-based government can be undertaken, and a model of how to integrate policy development, policy implementation and policy evaluation from the outset and over the life course of a policy.

### **Conclusion**

Evidence-based government and evidence-based policy have become familiar terms in many countries, so much so that one wonders whether they are sometimes little more than a rhetorical device. This paper has argued that in the UK context, at least, evidence-based policy and evidence-based government is well established. However, a broader conception of evidence is used by most governments than that used by some academics, and a range of methods for gathering and appraising evidence for government purposes is quite common.

The driving force for evidence in government tends to be the type of question being asked, rather than any particular research method or design. Most governments require sound evidence on both the effectiveness of outcomes and the effectiveness of implementation and delivery of policies, programmes and projects. The availability of both types of evidence is often in short supply. The challenge of this for groups such as the Campbell Collaboration is quite clear: there is much to be

done and the delivery of useable products is overdue.

This paper has also argued that there are influences on government and policy making other than evidence. These include the experience, expertise and judgement of policy officials and Ministers, values and ideology, available resources, habits and tradition, lobbyists, pressure groups and the media, and the pragmatics and contingencies of everyday political life. The uncertainty of social, economic and political events, coupled with the uncertainty and inconclusive nature of scientific knowledge, mean that high quality evidence often has to compete with these other factors as a basis for decision making. Failure to appreciate these realities of government may result in the principles of evidence-based policy and practice being used less often and with less seriousness than they deserve.

## References

- Britten N, Campbell R, Pope C, Donovan J, Morgan M, Pill R., 2003.  
'Synthesis of qualitative research: a worked example using meta ethnography',  
Journal of Health Services Research and Policy, **56**, 4, 671-684.
- Cabinet Office 1999a  
Modernising Government, White Paper, London, Cabinet Office.
- Cabinet Office, 1999b  
Professional Policy Making for the Twenty-First Century, Cabinet Office Strategic  
Policy Making Team, London, Cabinet Office.
- Cabinet Office, 2000  
Adding It Up, London, Performance and Innovation Unit, Cabinet Office.
- Cabinet Office 2001a  
Better Policy Making, Centre for Management and Policy Studies, London, Cabinet  
Office.
- Cabinet Office, 2001b,  
Regulatory Impact Appraisal, Regulatory Impact Unit, London, Cabinet Office.
- Cabinet Office 2003a  
The Magenta Book, Government Chief Social Researcher's Office, London,  
Cabinet Office. Available at <http://www.policyhub.gov.uk>.
- Cabinet Office, 2003b  
Trying It Out: The Role of 'Pilots' in Policy-Making, London, Government Chief  
Social Researcher's Office, London, Cabinet Office.
- Cabinet Office, 2003c  
The Design of the Employment Retention and Advancement Demonstration Project,  
London, Government Chief Social Researcher's Office, London, Cabinet Office.
- Campbell R, Pound P, Pope C, Britten N, Pill R, Morgan M, Donovan J., 2002  
Evaluating meta-ethnography: a synthesis of qualitative research on lay experiences  
of diabetes and diabetes care. Social Science and Medicine, **7**, 4, 209-215.
- Chen, H.T., 1990  
Theory-Driven Evaluations, Newbury Park, Sage Publications.
- Cooper, H. and Hedges, L.V. (eds), 1994  
The Handbook of Research Synthesis, New York, Russell Sage Foundation.

- Davies, P.T. 1999a  
 'What is Evidence-Based Education?', British Journal of Educational Studies, **47**, 2, 108-121.
- Davies, P.T. 1999b,  
 'Introducing and Monitoring Change' in Dawes, M.G., Davies, P.T., Gray, A., Mant, J. and Seers, K. (eds), Evidence-Based Practice: A Primer for Health Professionals, London, Churchill Livingstone, 1<sup>st</sup> Edition.
- Davies P.T. and Boruch, R. 2001  
 'The Campbell Collaboration', British Medical Journal, **323**, 294-295.
- Davies, P.T. 2004  
 'Changing Policy and Practice' in in Dawes, M.G., Davies, P.T., Gray, A., Mant, J. and Seers, K. (eds), Evidence-Based Practice: A Primer for Health Professionals, London, Churchill Livingstone, 2<sup>nd</sup> Edition (in press)
- Davies, H.T.O., Nutley, S.M. and Smith, P.C., 2000  
What Works? : Evidence-Based Policy and Practice in Public Services, Bristol, The Policy Press.
- Davies, P.T. 2003  
 'Systematic Reviews: How Are They Different From What We Already Do?', in Anderson, L. and Bennett, N. (eds), Developing Educational Leadership for Policy and Practice, London, Sage Publications.
- Dixon-Woods M, Fitzpatrick R, Roberts K., 2001  
 'Including Qualitative Research In Systematic Reviews: Problems and Opportunities'. Journal of Evaluation in Clinical Practice, **7**, 125-133
- Dowd, S.B, 1994  
Education as a strategy for allied health, Technical Report, Chicago Illinois, Lincoln Land Community College.
- Gray, J.A.M. 1997  
Evidence-Based Healthcare: How to Make Health Policy amd Management Decisions, New York, Edinburgh, London, Churchill Livingstone.
- Greene, J.C., Benjamin, L., and Goodyear, L., 2001  
 'The Merits of Mixed Methods in Evaluation', Evaluation, **7**, 1, 25-44.
- Grimshaw, J.M., Thomas, R.E., MacLennan, G., Fraser, C., and Ramsay, C.R., 2003  
 'Effectiveness and Efficiency of Guideline Dissemination and Implementation Strategies', Final Report, Aberdeen, Health Services Research Unit.

Harden, A., Oliver, S., Rees, R., Shepherd, J., Brunton, G., Garcia, J., and Oakley, A., 2003

'An Emerging Framework for Synthesising the Findings of Different Types of Research in Systematic Reviews for Public Policy', Paper presented at the 3<sup>rd</sup> Annual Campbell Colloquium, Stockholm, Sweden (available at <http://campbellcollaboration.org>).

HM Treasury, 2003

The Green Book: A Guide to Appraisal and Evaluation, London, HM Treasury.

Langewiesche, W., 2004

'The Problem with Houston', Observer Magazine, 11 January 2004, 22-33.

Lomas, J., 1993

'Retailing research: increasing the role of evidence in clinical service for childbirth', The Millbank Quarterly, **71** (3), 439-475.

Martin, E.M. 1997

'Conditions that facilitate the school restructuring process', paper presented at the Southern Educational Research Association meeting, Austin, Texas, January 1997.

Mulgan, G., 2003

'Government, Knowledge and the Business of Policy Making', Paper given at the Australian National University, Conference of the Australian Public Policy Network, 23 April, 2003, Canberra.

Nutley, S.M. and Davies, H.T.O., 1999

'Introducing Evidence-Based Practice: The Contrasting Experiences of Health Care and Probation Service Organisations in the UK', Public Policy and Administration, **14**, 4, 39-57.

Nutley, S.M. Walter, I., and Bland, N. 2002,

'The Institutional Arrangements for Connecting Evidence and Policy: The Case of Drug Misuse', Public Policy and Administration, **17**, 3, 76-94.

Nutley, S.M. Walter, I. and Davies, H.T.O. 2003

'From Knowing to Doing: A Framework for Understanding the Evidence-into-Practice Agenda', Evaluation, **9**, 2, 125-148.

ODPM, 2000

Integrated Policy Appraisal, London, Office of the Deputy Prime Minister.

Office of Science and Technology, 2001

A Code of Practice for Scientific Advisory Committees, Office of Science and Technology, London, Department of Trade and Industry.

Phillips Inquiry, 2001

The BSE Inquiry: The Inquiry Into BSE and Variant CJD in the United Kingdom, London, Ministry of Agriculture, Fisheries and Food (MAFF). Available at [www.bseinquiry.gov.uk](http://www.bseinquiry.gov.uk).

Rawls, J. 1972

Theory of Justice, Oxford, Clarendon Press.

Rogers, P.T., Hacsı, T.A., Petrosino, A., and Huebner, T.A. (eds), 2000,

'Program Theory in Evaluation: Challenges and Opportunities', New Directions for Evaluation, No 87, San Francisco, Jossey-Bass.

Royal Society, 2002

Foot and Mouth Disease 2001: Lessons To Be Learned Inquiry, London, Royal Society. Available at [www.royalsoc.ac.uk/inquiry](http://www.royalsoc.ac.uk/inquiry).

Spencer, L., Ritchie, J., Lewis, J., and Dillon, L., 2003

Quality in Qualitative Evaluation: A Framework for Assessing Research Evidence', Government Chief Social Researcher's Office, London, Cabinet Office

Stocking, B., 1992

'Promoting Change in Medical Care', Quality in Health Care, **1**, 56-60.

Weiss, C.H. 1998

Evaluation, Second Edition, New Jersey, Prentice Hall.

Williams, T., Hill, M., and Davies, R., 1999

Attitudes to the Welfare State and the Response to Reform, Research Report No. 88, London, Department for Work and Pensions.