Evidence-based policy at the Cabinet Office

Phil Davies – Impact and Insight Series – ODI – 17 October 2005

I was asked to talk about evidence-based policy at the Cabinet Office. When I looked at this last week, I thought, it could be a very thin talk if I only talked about the Cabinet Office. Not a lot happens there in terms of getting evidence into policy except for the Government and Social Research Unit's job of analysing policy, and the work of the Prime Minister's Strategy Unit which, as it says on the tin, is trying to think more strategically about policy in five, ten and fifteen year cycles. The Strategy Unit should, therefore, be more evidence-based than other areas of government - or at least as evidence-based. So since I thought the title 'evidence-based policy at the Cabinet Office' may not be too fulfilling for you, I decided to change it to 'making public policy more evidence-based'. That seems to be what we are interested in around this table, and it is certainly what we are trying to do. It means identifying the opportunities for doing this, as well as the barriers and how to overcome them.

I will give you a brief outline of what I want to talk about. Firstly, I want to give you the reasons why I think we need evidence-based policy. Secondly, I want to talk about factors other than evidence. Appreciating the complexity of the things which compete with evidence is one of the most important things I have had to learn since transferring from academic life to working in government. (I have been working in government for five years now.) I also want to talk a little bit about different kinds of evidence because there is a huge contested field about what counts as evidence, particularly research evidence. Lastly, I have identified four major problems and I will offer some solutions.

The first reason why I think we should be interested in evidence-based policy is above all for improving effectiveness. It would be very good if in government we did more good than harm. I take that to be fairly self-evident, but there is sometimes an assumption in government circles that policy is about absolute good - and that any policy idea you came up with must be good because you came up with it. As we know that is not always all the case. We can all cite chapter and verse about policies which have gone disastrously wrong, and just the notion that we not only may be ineffective, but may actually cause more harm than good, is an important thing for us to consider. This is a concept which I have borrowed from twenty years in evidence-based medicine. We discovered in the 1990s that medicine could often do more harm than good and often the best way to intervene was by doing nothing.

The second reason is efficiency. We are all using scarce public resources and therefore trying to get the maximum effect from them. I will come back to that in a moment. Certainly, within the rhetoric of the Blair administration, we are supposed to have a service orientation: we should be meeting citizens' needs and expectations. Indeed, something which we are working on at the moment is a customer satisfaction index for the whole of government, which is quite interesting (though I will not go into it here).

There is also an accountability issue. I think that accountability is one of the major advantages of having a more evidence-based approach. Whatever the area of policy, whether it is medicine, education, international justice, social welfare or international development, it is very nice to know what people do and why they do it. One of the major developments in this respect has been the Freedom of Information Act. Our department gets many enquiries from the public wanting to know why the government has done x, y or z. It has been really positive to see the level of public involvement over the first nine or ten months. I am sure that individual departments get enquiries as well. It is important that people are aware of what government is doing and of the effectiveness and efficiency criteria. This is an important part of the democratic process. Given that we are trying to enhance or restore more trust in government (which seems to be yet again in doubt), I think that this can enhance the establishment of trust by making government more open, transparent, service-oriented and efficient.

To me, evidence-based policy is in essence quite straightforward and the driving reason behind why we want evidence-based policy is that it is about helping people to make well informed decisions by putting the best available evidence from research at the heart of development and implementation. It does not matter whether the decisions are in government, in practice, or any other decision-making context, nor whether they are at the local, national, or supranational levels. The two words which are contentious within that definition are 'from research'. The problem I think we have is that people want to use evidence, but they do not necessarily want to use research evidence. That has been one of my problems working in government. Research evidence is not as valued as much as perhaps some of us around this table would like it to be.

I want to consider some of the other factors that influence policy making. When research evidence is not a factor, what else is driving policy? Governments, departments and the civil service are like many other organisations. They are full of people who claim - and indeed may have - expertise and experience which they often feel trumps the evidence. At one level I think that this is right - it is important not to disregard the experience which people have, whether that is substantive expertise or the kind of tacit knowledge which many people believe is at the heart of good practice. This kind of knowledge is about knowing when to do things and what to do and it does come with experience. Similarly, judgement is very important. It is what politics and good decision-making are about and skills of good judgement are developed over time. But it would be nice to see that judgement based not just on expertise but also on research evidence.

The issue of efficiency is also at the heart of another very real problem which we are likely to be facing over the next few years, which is the problem of declining resources. It has been a relatively good time since 1997 and there has been a pretty much year-on-year increase in public expenditure in the areas of interest to those around this table. It will probably not be like that for a few more years given the slowing down of the economy. As we know, in the 1980s and 1990s, the availability of resources to many of the programmes that we are interested in was very sharply defined.

When people say that evidence-based policy is about what works, I am not sure that is quite true. It is about what works, at what cost and with what outcomes. It would be extremely myopic, if not stupid, not to realise that politics is also about values and about political contexts. There is an excellent piece in the Journal of International Development, which John Young and others shared with me, in which this point is made very clearly. The policy context, and indeed the political context, is highly important to understanding when and how you can use evidence. John's own paper was talking about international development, but I would say that it is just as important in the domestic context. We need to acknowledge that people have fought pretty hard to get into power, based largely on values and beliefs and self-beliefs. I would put these near the top of the list of factors which are competing with evidence and with trying to get people to think in a more evidence-based way.

Habit and tradition are things which most organisations have to deal with, but believe me, in the civil service, parliament and government, there is a very habitual and traditional way of doing things. When we sometimes ask why we are doing things in the Cabinet Office, the answer is 'because we have always done them that way'. Very often that is how organisations function. One of the tasks of any evidence-based practitioner should be to challenge that and to suggest another way of doing it. I worked as a lecturer in a university for twenty years which, as many people here will know, is an extremely habitual and tradition based university. With two new vice chancellors in recent years they are starting to think about things in a different way.

Lobbyists and pressure groups play a crucial role. I am beginning to think that if, as evidence-based policy practitioners, we want to get things done, we ought to become a lobby group or a pressure group, or certainly to learn some of their tricks. I am amazed at how good they are at getting very simple messages into practice very quickly, and we are not. So I am trying to learn more about how it works.

My final factor is sheer pragmatics and contingencies. Things just happen and when something like Foot and Mouth or bird flu breaks out, we find that no one had really thought about it. People start doing unproven things, or doing nothing at all. Nobody really imagined that two planes would fly into the side of a building in New York, or that people will put bombs on the Underground, but once that happens, the whole prioritisation and modus operandi of government changes and you have to be very careful not to loose track of what the government was trying to do. This is true for other organisations as well; most organisations had to change dramatically in light of those major national incidents. The Iraq war happened and as a consequence of that we now have yet another fuel crisis, which in my estimation is likely to get worse before it gets better. That undermines the whole basis on which the overall five or ten year programme of government was built. Whilst every government has contingencies, it is very easy to go outside those contingencies and to have to think again.

My view is that it is when these things happen that is when we should turn to evidence. Most people in government say that that is when you cannot use evidence; you have to be a decision-maker, use judgement, expertise. That is true, but if you think of two of the highly chronicled disasters of policy in the last ten years, BSE and Foot and Mouth - BSE was independently reviewed by the Phillips enquiry and Foot and Mouth by the Royal Society - and they made almost identical critiques of what went wrong. Both of them said that we should have looked at the evidence in a more systematic and analytical way, rather than trying to 'wing it' or listen to pressure groups such as the National Farmers' Union or whomever was advising (in that case wholly inappropriately). Rather than abandoning evidence, these are precisely the situations where you need to have evidence most.

This leads to the issue about what types of evidence we need for policy and practice and what types of evidence we actually use in government. A lot of government is, or should be, about trying to establish the impact of something - and whether the government's policy is going to have an impact over and above doing something else, or doing nothing at all. This goes back to my point about trying to do more harm than good. I passionately believe that in order to do that, we need experimental or quasi-experimental studies, where these can be done appropriately. I am talking about studies which include a counterfactual comparator and which can tell us what would happen if we did something else, or did nothing.

However, a whole lot of government is not just about determining what the impact of a policy would be, it is also about how to implement it and how to get it to work on the ground in very varied circumstances. Most of you here today work in the international arena, but even within Britain we often find that the implementation strategies which we need for a welfare to work policy, or a health or education policy, may be massively variable across the country. We have very varied demography, labour markets and culture across Britain. For this, again, I think we need experimental and quasi-experimental research, but we also tend to need much more qualitative and descriptive work, interview work and focus group work. We certainly need ethnographic work, and we do not use good ethnographic work in government nearly as much as we should. We spend a lot of money on qualitative research, possibly around half of the total amount we spend on external research in a year, and I have to say I am less than impressed with the quality of much of it. I am also less than impressed with what is done with it, but the two are no doubt somewhat related.

I also think that we need to be very clear about what the theory of change is that is that we are trying to use and that is being pursued between policy inputs and policy outcomes. This is my fifth problem: that policy-makers tend not to think in terms of the theory of change. The very word theory often closes policy-makers down and they do not want to know, so I often use the term 'logic model' or some sequential model. How do you expect making children do homework for two hours a night to raise literacy standards? That is government policy, but what is the logic by which we ask children to do homework? We need to unpack the logic by which we are asking kids to do homework, particularly when it is not always clear that it works. Of course if you ask some children to do their homework in twenty minutes they will do it all absolutely fine, but other children will take a lot longer and need a lot more help with it. We need to unpack the theory of change at that level, and at the outset of making policy.

We do need descriptive analytical evidence and we utilise it a great deal in government. The government sponsors dozens of major surveys and are now pushing for one much more coordinated survey which will bring together a number of other government surveys. The government owns acres of administrative data and it is nice to be able to view this in a comparative context. Once every year or two years the Prime Minister's Strategy Unit does a strategic audit for Number Ten, where we look at the range of things going on in British society and compare them with other countries of comparable economic development.

I think we also need attitudinal evidence, by which I mean gauging roughly what the limits to government are, how far governments can go with public policy and how far they can take the electorate. You can always challenge the electorate and try to go beyond that, but if you do not even know 'where the electorate are at', you can find yourself in dire trouble, as I believe the Thatcher administration discovered over the Poll Tax. The Poll Tax was clearly never going to work from day one, partly because they had piloted it in Scotland and it had not worked there and partly because of historical reasons, but it was quite clear that it simply did not go down well with ordinary voters. They could see through it very easily.

We do need to know a bit about numbers, what would happen if we did x and y and z. We need scenario planning, statistical models and various aspects of multivariate analysis. Because of the resource issue which I mentioned earlier and concerns about efficiency, we do need cost-benefit data, cost-effectiveness data and cost-utility data. We also need at least a few people who are able to do and understand the econometrics of that and use some of the statistical modelling techniques to see what sorts of outcomes we are going to get with different scenarios for public expenditure.

My last point about the types of evidence we need relates to a broader issue to which we give far too little attention in government and that is ethical evidence. We need to consider in ethical terms the rights and wrongs, pros and cons, and advantages and disadvantages of doing various things. Given that most of the task of government is about allocating scarce resources and making trade-offs, there is a question about the basis on which those trade-offs are made - not just a technical or technocratic basis, but an ethical one. Whether we are looking at decisions concerning road versus railways versus other modes of transport, or whether you are looking at the care of the elderly or the care of very young children, we are invoking certain models of how we value life and resources. This begs the question of who is making those decisions. There have been a lot of attempts in various areas of public policy to carry out public consultation, such as the Oregon method in United States' healthcare, the town hall community methods, citizens' juries that have been tried in Britain, and various public consultation mechanisms. My point is that these decisions and issues are not just for decision-makers in Whitehall or in local authorities, PCTs [Primary Care Trusts] or education authorities to make; they are issues we should be addressing in a more collective way. I am not sure of the best ways to do that, but I do think it is important that we take into consideration ethical evidence.

I want to avoid the notion that all you need to have is a set of randomised control trials and the world will be a better place, or the idea that if we had the most definitive ethnography of how to make Easterhouse a nice place to live in we would be alright. I think that those single discipline, single methodology arguments are pointless, erroneous and do not help us to build the balanced evidence-base that we need for decision-making. You have to decide in each case what types of evidence you need for each type of question.

We have five problems in terms of how we get from opinion-based models to evidence-based models of policy making. First of all, social scientific knowledge is often not up to the challenge. Not all the social scientific research that is done is of sufficient quality. In doing quite a lot of systematic review work, the specific problems I have found with it are: unclear objectives, poor research designs, bad or selective data reporting, methodological weaknesses, unsupported conclusions and the sheer uncertainty of unscientific knowledge. In some areas of public policy, the political and social scientific knowledge is very uncertain. In some areas of government you get moves towards evidence-based policy. The Royal Society's review of the Foot and Mouth outbreak of 2001 showed that some of the policies of the Department for the Environment, Food and Rural Affairs (DEFRA had an evidence base something like this [see graph]. DEFRA is now building up its social and economic capacity and improving the evidence-base of its policies There may be other departments trying to develop an evidence-base, such as the Department of Education and Skills trying to find out the most effective ways to teach mathematics and science. We know a lot more than we did ten years ago but there are still a lot of areas which remain highly contested and uncertain. On the other hand some of the Welfare to Work and the New Deal policies which have been introduced by DWP started from a high evidence-base from the outset which left little scope for opinion-based policy. These scenarios can be built both within and between departments. We have a very variable social scientific data and knowledge base.

The second problem is that there are often very different notions of what counts as evidence between policy-makers and researchers. These findings [see slide 11] come from Jonathan Lomas' work in Canada, but I will also report to you the findings of some work which we have just finished on the Analysis for Policy project which has come to almost identical conclusions. Lomas found that policy-makers are quite prepared to look at evidence which is colloquial and highly contextual - sometimes referred to as anecdotal evidence. They will accept anything which seems reasonable. They want it to be policy-relevant and timely and they want a clear message with low contestation. In our study we have come to almost identical conclusions. Researchers' evidence tends to be obsessed with being scientific, context-free, rule-bound, empirically proven or demonstrated, theoretically driven and conducted within as long a timeframe as necessary. We social researchers tend to emphasise the caveats and qualifications, but often tend to forget that getting a clear message across is just as important. These are two very different worlds competing with each other. We are now trying to bridge these two communities and doing so is about knowledge translation.

In the work which we were doing on the Analysis for Policy project, we asked over fifty senior policy-makers across various government departments about the types of evidence which they used. They wanted quantitative and statistical evidence, economic evidence, qualitative and survey evidence. The most interesting finding was that a common type of evidence they used could be described as anecdotal evidence, including anything which could give them 'a finger to the wind' or a 'good intuitive instinct' to indicate whether they were heading in the right direction. They wanted both hard and soft evidence, from national and international sources. Two types of evidence which they did not mention were experimental evidence (which most had never heard of and did not want to know about) and systematic reviews of evidence. I mention that because in the Cabinet Office, we are currently putting a significant amount of resources into developing these two areas, only to discover that not only have most people not heard of them, but they possibly care about them even less.

If you ask policy-makers where they go for their evidence, they will tell you that they go first to their special advisors, then to people who are called experts (in whom I have little faith), then to think-tanks and opinion formers, lobbyists and professional associations, media, their constituents, consumers and various users of services and only then, if they bother, will they turn to academics and research evidence. This point was also made by an internal piece of research in the Department of Trade and Industry in which a survey of their decision-makers found that academic research was not even mentioned. Our job in the Government Social Research Unit is to try to get academic research much further up the evidence chain.

Another point, which I can only illustrate, is about the uncertain logic and theory of change. Some of you will know about the 'Scared Straight' programmes. These are juvenile awareness programmes which are supposed to reduce offending by exposing young would-be criminals to prison life. The theory is that this exposure will frighten them so that they will not commit crime. The programme evidence suggests that this does not work. Experimental evidence shows that kids who go through the programme not only do no better, but actually do a lot worse, and tend to commit crime much more than those who had not gone through the programme. The qualitative evidence which has been done on the programme is even more interesting, because it tells us why. It suggests that kids find prisoners to be positive role models of the people they want to be: tough, hard and the kind of guys that people are scared of. This in fact stimulates them to be criminals (and probably teaches them how to do it as well!). So it is not surprising that in the randomised control trials, the experimental group tended to go on to live lives of higher levels of crime. This research has been available for five years or more and has been accumulating for twenty years, and still people are using these programmes.

Another issue is the accessibility of research evidence. We asked those senior policy-makers who had used research evidence to tell us what they thought about its accessibility. They found research evidence too long, too verbose, too detailed, too dense, too impenetrable, too full of jargon, too methodological, too untimely and, above all, either non-relevant or irrelevant. These are their perceptions of what we provide in the social research industry. Again, we are going to do some follow-up work to test this out in a more survey-based way across government, as these findings were based on qualitative interviews with fifty senior policy-makers. If this turns out to be the case, you can see some of the problems which we are facing.

We are working on a number of solutions. Firstly, we are trying to integrate research planning into policy planning across government in a strategic way. CRAG stands for the Coordination of Research and Analysis for Government. It is a fairly high-level initiative which seeks to bring all government departments together at the level of senior civil servants (which includes chief scientists, chief social scientists, chief statisticians, chief economists and Permanent Secretaries) to try to plan policies across a five or ten year time-span and to plan the research that is need to support these policies. Above all, to make sure that the policies and evidence are joined up and coherent.

Secondly, the PSG (Professional Skills in Government) initiative is now giving policy-makers some incentives for evidence-based policy by requiring them to demonstrate in their annual performance reviews that they can and have used evidence. This means that if they cannot do this, they may forfeit any personal bonuses. The idea is to give people actual incentives to get involved in using evidence, including taking part in some of the training initiatives that have been established.

We need to establish the signed-up ownership of the evidence by the senior policy-makers who often commission research. When these people move on, quite often no one wants to know about the research that has been commissioned. When we are spending half a million pounds (or even lesser amounts) on research, we cannot afford simply to walk away saying that it was the interest of the previous minister or official. It is unacceptable.

Researchers need to qualify these policy and practice issues at the outset. What is it that policy-makers want to know? What is their theory of change? What evidence is important for us to gather. This means establishing a theory of change or a logic model upfront and trying to get answerable questions. Questions such as 'do prisons work?' are not answerable questions, they are too simple questions. A question which we were asked recently by someone in government was about what evidence we had on the relationship between globalisation and crime - and whether they could have it by four o'clock! We asked more about what they actually wanted to know and it turned out that what they really wanted to know was how use of the internet had increased credit card fraud, which is an answerable question. We could find out the answer to that in two or three days, but the first question was an unanswerable question.

We in the research community must understand the policy timetable. Colleagues in the University of York recently suggested that I should try to slow down the policy process, but my job is to work to it and not against it. We need knowledge translation and to find a signal through the noise of research, even if that is a weak signal. Now that I have turned from poacher to game keeperr, I get sent a lot of reports from across government, many of which I haven't a clue what they are about. I should not have to dig for four days to find out what a report which has cost us a,lot of money is telling us. Various formats for good reporting are available and there are people around this table who have devised them. I like the Canadian Health Services Research Foundation format, which requires that every report which comes to government must have a one page bullet-pointed message, a three page executive summary and a final report no longer than twenty-five pages - if only because no one ever reads reports longer than twenty-five pages in government. If we want to write a 250 page book about something, that is fair enough, and it is good to have a copy and know that someone has done that work, but I can assure you that no one in government will read it.

Lastly, I want to make a point which I have borrowed from Matthew Taylor, who spoke at one of your earlier seminars. Matthew makes most of the points which I have made above, but one argument which I did not think of is one which came from IPPR (the Institute for Public Policy Research) and from that think-tank mentality. The argument is that whatever your message, if you want to convey it, you have to be consistent and to keep hammering on about it time and time again - as indeed Labour did just to get into power. Their strategy was to always keep giving the same message (even if it was the wrong one) and never go off-message.

The other side of that is about being opportunistic. The Campbell Collection's systematic review has just been published on counter-terrorism. It could not have come at a better time. After the July 7th bombings, suddenly the whole of government wants to know what to do about terrorism. On my desk lands the first review that I have seen from Campbell in a year and it goes straight to the heart of what we need to know. But it also took me two days to decipher the key findings and when I checked it with the author she confirmed that the basic conclusion was that, when it comes to counter-terrorism, not a lot works. Nonetheless, the report is now with various decision makers in government and this is probably the first time a Campbell review has ever been read by someone outside the Campbell Collaboration, and it was entirely opportunistic.

Comments and questions from the floor included:

- To the question on whether beliefs can be considered evidence, Phil Davies responded that it is important for decision makers to know what the public's beliefs are in other to understand the context in which they are making policy.
- Instead of talking about 'Theories of Change' to describe the change narratives that policymakers have about the solutions to particular problems, it was suggested to call them 'Lines of Argument'.
- It was agreed that it is also very important to consider the demand for evidence: what research is actually needed.
- Although the presentation focused on the policymaker's perspective, it was considered important to think about the incentives for researchers to develop policy relevant research
- It was emphasised that researchers need to make sure that they make their evidence more accessible. Often, academic research does not meet the requirements satisfy the demand for evidence.