Evidence-based politics. Government and the production of policy research.

The LSE GV314 Group

Introduction
Governments continually affirm their belief in evidence-based policy making. The faith in the ability of scientists, natural and social alike, to help policy makers only occasionally achieves the kind of rhetorical prominence it did in the early years of new Labour after 1997 or with Harold Wilson in 1963, but it can always be detected to a degree in government thinking and actions over at least the past half century. Yet the current Conservative-Liberal Democrat coalition has the strong reputation among many researchers of having turned its back on "evidence-based policy". There seems to be less policy research commissioned since 2010 and some ministers have put their disdain for research on public record (see Levitas 2012); Eric Pickles (2011), for example, was reported as saying that as far as evaluation studies are concerned "in my experience "you can probably get an academic to do anything you want".

It is too soon to determine whether any such drop in enthusiasm for using social research in policy making is permanent or temporary. Tarling (2011) suggests there has been a longer-term decline in trust between academics and policy makers. There are signs, however, of the enduring strength of the appeal, whether substantive or rhetorical, of evidence-based policy making. Government departments still commission policy research, albeit with less frequency than under the preceding Labour government. The Government Office for Science (2012: 20) comments on the Department for Work and Pensions' "strong commitment" to using analytical and scientific evidence to inform the development and the delivery of policy.

The idea that evidence is "embedded in the political and policy rhetoric" of the twenty-first century and "infused in the newly-transformed professional ethic of many service professionals" (Davies, Nutley and Smith 2000:11) is all the more puzzling if one considers that most empirical studies suggest there is little real chance of government using research to shape policy in a direct way. The idea that governments base decisions on whether to extend, amend or terminate a programme on evaluations of them flies in the face of most of the evidence we have about research utilization, even research conducted in what appears to be the most favourable conditions for its utilization. "Most studies seem to be used in selective bits, reinterpreted to fit existing preferences or ignored" (Weiss et al 2008: 30).

To judge evaluation research by its direct impact on policy making almost invariably finds its role in policy making to be small. A range of scholars seeking to assess the impact of social scientific policy evaluations on policy tend instead to emphasise the less direct impacts of adding to cumulative knowledge and understanding about the characteristics of policy interventions that can at some unspecified time be brought into the policy making process (Weiss 1977; 1995; for a review see Weible 2008). Research evidence does not have a direct "instrumental" use but can have less direct "conceptual"
and "symbolic" uses in policy making (see Davies 2012). The literature on research utilization contains many convincing accounts of why research does not have much of a direct influence on policy (see Beyer and Trice 1982 for a meta-analysis), including those based on differences in timescales (Jowell 2003: 9-10), professional environments (Martin, Currie and Lockett 2011) and modes of argumentation between the worlds of science and politics (Ritter 2009) on and institutional constraints on policymaking (Waddell et al 1995), such that this lack of direct or "instrumental" influence can be described as overdetermined.

Why policy research continues to be attractive to government and researchers despite its known lack of such direct influence might not be hard to explain. For the people that write it there is a range of possible employment-based benefits: if they are in-house researchers it is part of their job, if they are externally contracted then they and/or the organisation for which they work are paid for it. Moreover, working on contract for government can bring benefits in terms of status, fulfilment and furtherance of a research career. For the people that commission research, there is a range of possible political benefits: politicians can expect their policy choices to be given the patina of scientific validation. Commissioned research may be an effective way to advertise government concern and activity. In addition, policy makers, officials and service providers might hope for a science-based vote of confidence in their programmes and strengthen its chances in the competition for resources in government.

Government commissioned policy research cannot simply follow the simple principle of "who pays the piper calls the tune" and impose political objectives on the hired hands that work for it because doing so would not produce anything recognized as scientific research and thus undermine both the political and employment benefits. To serve its employment and political objectives, policy evaluation research must to a significant degree follow what might be termed "evidence-based policy" objectives: the development of insight into policy alternatives through the rigorous analysis of causes and/or effects based on the application of scientific techniques and methodologies.

The production of government commissioned research is thus likely to involve some degree of compromise between employment-based, political and evidence-based policy objectives. Our concern in this paper is in the tradeoff between political and evidence-based policy objectives since these two sets of objectives get at the heart of the value and credibility of government commissioned research. Our examination of this trade-off focuses on exploring how political objectives interact with evidence-based objectives in the progress of commissioned research.

While it might at first appear desirable to come up with a clear answer to the question of whether political objectives have too much (or too little) role in commissioned research, the quest for an answer to this question is unlikely to bear fruit. We have no clear metric by which to judge sizes of roles, such judgments are likely to be contested and contestable and such roles are contingent on too many variables (ranging from the strength and ideology of the government to the characteristics of the researchers) to be susceptible to an easily generaliseable answer. Instead we focus on understanding how
common it is for governments to pursue political objectives in commissioning, managing and using research, how such objectives are pursued and the reactions of the researchers conducting the research to the pursuit of these objectives. Our analysis concentrates on the kind of research where one would expect the potential for conflict between the two sets of objectives to be highest: in research evaluating the impact and/or success of government policies.

Where politics meets science
Some conflict between political and evidence-based objectives might be expected to be found at all stages of the research process. In selecting the topics for scientific evaluation we would expect evaluations, if they were to follow an evidence-based logic, to be conducted on issues where there is a prospect that their results could matter. This might mean that the policy is significant -- perhaps because of its actual or potential importance for a large number of citizens (e.g. workfare programmes) or for a cherished principle that affects few citizens directly (e.g. foreign aid) -- and that there are at least some aspects of this policy that might be open to change. A political objective for selecting a topic for evaluation would be in the potential for making the government look good and its opponents bad. This might be the case where government commissions research on a topic that it has no ability, resources or intention to address, simply to indicate concern with it; or where government expects research to endorse policies or otherwise give public credit to (or remove blame from) those who devised the policy (see, for example, Ham, Hunter and Robinson 1995: 71).

In managing the scientific inquiry one would expect a government pursuing evidence-based objectives to orient its continuing relationship with the scientific researchers to the provision of advice and guidance in maintaining the policy focus of the research concerned and providing data and support in gathering evidence. A political logic of management would reflect the desire to ensure that the results benefit the government politically, above all by trying to steer the research so that it reaches desired conclusions, usually positive (or at least not negative) about the policy. In utilizing the results of the commissioned research, an evidence-based logic does not necessarily mean that results will be instantly acted upon. In fact, we know that it is very difficult for evaluations to have a significant direct effect on policies even with, apparently, the best intentions and efforts of policy makers and the best conditions likely to be available (Weiss 1990; Greenberg and Mandell 2007). Such evidence is at best but one potential source of policy advice anyway. The difference between an evidence-based and a political approach to the use of policy research is primarily in the way in which policy makers bring scientific evidence into the mix of considerations that help shape policy. An evidence-based logic suggests that all relevant evidence systematically developed has value and plays a valid role in shaping policy; a political logic suggests that only the

\[1\] While it is possible that political objectives behind commissioning an evaluation might include producing a report critical of a policy -- possibly because the person/people commissioning it would like to see the policy discontinued or its champion undermined or discredited -- we assume in this paper that political objectives are more frequently served by producing positive findings.
evidence supportive of the values of the politician commissioning the research plays a role in shaping policy while that which is not supportive is ignored.

The study
To explore the question of how these two potentially conflicting objectives of research play out in British government we conducted a survey of academic researchers who have been involved in commissioned research. To gain our sample we looked at 215 research reports published since 2005 and noted the names of all the authors of the reports, where named. These 215 reports were not gathered through random sampling but were all the reports from 2005 onwards made available online in autumn 2011 on government department websites. Report selection and consequently selection of our sample is thus vulnerable to biases reflecting the publication and archiving practices of individual websites. We do not know what these biases are. We have no particular reason to suspect any systematic bias in what reports are put on the website and for how long they remain there, but since this appeared to be the only available method of developing a sample of researchers we had no choice anyway. Our study covers academic researchers only, that is to say, those working for a university or an organisation based in a university. We chose academic researchers for practical reasons. In our early approaches to researchers none of the researchers working for non-university organizations we initially contacted (by email) replied; almost all academics replied. At this early stage we interviewed 22 academic researchers.

Of the 215 reports we looked at, clear authorship information was given for 204. Of the 204 reports 59 were written by academics alone and a further 36 by academics working alongside non-academics. These 95 reports involving academics (written for 11 departments) yielded 275 unique names of academics for whom we could find email addresses. We supplemented this list with an additional 106 names of academic researchers whose university websites indicated they had worked on government commissioned evaluation research since 2005. We sent out 381 requests to complete our survey which was administered by Bristol Online Surveys in February and March 2012. 24 of our email requests were returned undelivered. After two reminders we received 204 questionnaire responses from 357 emails that were successfully sent, indicating a response rate of 57 per cent. We were particularly fortunate that exactly half of our respondents chose to write comments in the space we left for the purpose at the end of the questionnaire, some rather lengthy. The average number of words written by the 102 who used the space was 94 (these will be referred to as "free comments" when quoted).

The variability of the government research environment
Our research can only indicate the broad trends in the interaction between the political and evidence based logics in the conduct of commissioned research. It is certainly possible that different parts of any one piece of research display different balances between the two objectives. It might, for instance, be expected that politicians show greater interest in shaping the more controversial or politically salient aspects of a piece of research while leaving the researchers free to develop other parts as they see fit. We would certainly not expect our research to allow us to characterise any single piece of research as supporting evidence-based or political logic exclusively or even
predominantly, but rather the general consequences of the interaction between the two.

The type of interaction might, as our interviewees as well as our questionnaire's free comments suggested, vary according to the particular government department concerned. While such variation is likely to occur along a variety of dimensions (e.g. on different dimensions relating to how different departments go about commissioning research, how they manage it and so on), one free comment gets across the broad importance that many respondents and interviewees attach to departmental variation:

Government departments vary considerably … I used 'neither agree nor disagree' [in answer to one question on the questionnaire] to reflect the fact that I have very good experiences of the Department of Health and the Department for Work and Pensions in commissioning research but much poorer experiences with respect to Home Office/Ministry of Justice and the Department for Education (the latter by repute - I wouldn't touch them with a bargepole personally).

We did not, however, ask questions which would allow us to compare experience of different departments. The questionnaire was completed under guarantee of anonymity. The larger part of the questionnaire asked questions concerning the respondent's most recent research project. We did not ask which department commissioned it because when the question was included in a pretest we suspected it might be responsible for a low rate of response. Moreover, the population was not large enough to expect such a question to be of value.

Respondents further indicated in free responses and interviews that the research environment varied intra-departmentally. Among the commonly mentioned sources of variability were the time pressures on the project, the priorities and interests of the minister, the party in government, the period in the electoral cycle and the character of the research supervision in the department. While some of these points will be taken up below, we do not examine these as factors shaping the progress of research. Our central concern in this article is with understanding the overall pattern of accommodation between political and evidence-based objectives in commissioned research rather than explaining variability in why any single pieces of research were or were not successful, influential or otherwise satisfactory.

Does government play it safe with commissioning?
It is difficult to offer a clear picture of the degree to which government evaluation

---

Naming a department along with the year the research was published as well as the age, gender, disciplinary background and position in the research team might lead respondents to feel they were identifying themselves or their research teams and thus be reluctant to respond. Moreover, with around eleven departments covered in the survey we were unlikely to gather sufficient cases in each to make any reliable comparisons. We might have expected enough in the departments well represented in our sample, but this would have meant, effectively, we would only be likely to be able to offer plausible (but not statistically significant) data comparing at best three departments: DCLG, Education and Health. We decided to drop the question identifying the department
research tends to concentrate on "safe" topics that at least avoid the likelihood that the results will embarrass it. Our questionnaire responses at first glance support the "play it safe" argument, albeit with indirect evidence. Evaluation reports do tend to be supportive of government policy more often than they criticise it: 41 per cent of researchers said their most recent research offered support or strong for "the soundness of the general ideas underpinning the policy or programme"; 15 per cent said it offered challenge with 34 per cent "an even mix" of support and challenge for the policy ideas they evaluated (the remaining 10 per cent said their latest work for government neither supported nor challenged policy as it was not an evaluation). We asked an additional question about the degree to which the latest report respondents had produced was critical of the way the policy was being implemented. Here the researchers were slightly more likely to produce more critical comment on government activities: 27 per cent said they offered support or strong support for the "way the policy or programme was being implemented or administered", 23 per cent said they primarily offered a challenge while 36 per cent said they offered an "even mix" of support and challenge over implementation.

How far this tendency to support the policy but be slightly more critical of the implementation is down to selection of topics for commissioning and how far this reflects other influences (e.g. pressure from government or restraint on the part of researchers) cannot easily be assessed. According to the most direct evidence on selectivity in commissioning from the survey, government does appear in general play it safe: 31 per cent of respondents agreed with the statement that "Government is often prepared to take risks in commissioning research that might produce results that highlight shortcomings in policies and programmes" while 49 per cent disagreed and 20 per cent neither agreed nor disagreed (N=185). A larger proportion of respondents felt government played it safe.

Some of our free comments highlighted wider political goals in the selection of research topics. One respondent wrote that "Evaluation tends to be funded to support rhetorical assumptions, and findings are buried if the assumptions are not supported", and another "One study I worked on for the department of health was like being in an episode of 'Yes, Minister'. The conclusion I drew quite clearly was that the study had been commissioned entirely to avoid having to change anything. The main purpose was to slow down and sidetrack demand for change" and a third "the last research I did was commissioned to respond to a critical select committee report. By the time the research was commissioned, undertaken and reported on, the policy agenda had moved on, as had the original minister and civil servants, and there was little interest in the results".

However, in the selection of topics for commissioned research we should not necessarily conclude that playing safe is a particularly dominant feature of commissioning behaviour. While more respondents believed government played it safe than did not, the margin (at 49 per cent to 31 per cent) was not overwhelming. Moreover very few academics (10 per cent) "disagreed strongly" that government was prepared to take risks. This was much larger than the 2 per cent who "agreed strongly", but is still rather more suggestive of a clear tendency to caution in selecting topics than an attempt to cook the books. Many programmes, especially large ones, have to be evaluated as part of the deal (above
all with the Treasury) which led to their development. Moreover, as we shall see below, there are different players within government departments and those who often have most to do with setting up and commissioning research -- research administrators and officials in research departments -- might be expected to be more sympathetic to evidence-based than political objectives and these officials might be expected to play a role in the selection of topics for research.

Other aspects of the commissioning behaviour of government departments caution against the view that the quest for commissioning favourable evaluations is a dominant concern in deciding which policies to evaluate. If we look at the evidence of our 215 recent evaluation reports (from which our sample of academic respondents was largely derived) the evidence is not strong that government seeks to skew the research process in a way that produces good news by only commissioning research likely to yield it. Of course, to offer a satisfactory answer to our question we would need some tally of the total number of policies eligible for evaluation and compare it with what was actually evaluated. In the absence of any such statistics we must rely on more indirect measures.

Government, according to our coding of the 215 recent research reports, unsurprisingly tended to focus its attention on recent policy innovations: 27 per cent were evaluations of pilot projects and 44 per cent evaluated policies that had been in existence for less than three years, with a further 27 per cent being studies of older programmes or general surveys of policy areas. So insofar as taking risks is concerned, the evidence is that governments evaluate policies that are likely to reflect on their own performance, not that of a predecessor. Moreover government did not shy away from commissioning research that gave judgments on policy success. We examined whether policies were summative (i.e. formed judgements about the overall impact of research) or formative (looking at ways of improving the way the policy was being implemented or delivered). Only a minority of reports, 22 per cent, confined themselves to formative goals (sometimes termed a "process evaluation"), 46 per cent were summative and a further 30 per cent mixed summative and formative evaluations. This is broadly consistent with Salisbury et al’s findings (2010: 25-6), although it is important to note their suggestion that intentions to provide summative evaluations may not be matched by the research design and form of evidence used to develop them, and thus the quality of evidence may prevent a study intending to be summative from adequately filling that objective. Our coding of the individual reports put the proportion of critical studies (11 per cent) somewhat lower than that reported by the academics in our survey (16 per cent) and the positive reports higher (54 per cent) than in our survey (41 per cent).

Were academics more likely to produce reports critical of government than anyone else? Existing evidence rather suggests so (see also Salisbury et al 2010: 109), but according to our analysis of the coded reports they are not much more likely to be critical, though the numbers involved are too small to be statistically significant. In our coded reports database academics (at 52 per cent) were marginally less likely to produce a positive report than average (54 per cent Table 1). While reports produced in house, i.e. by researchers employed by government organisations, tend to be the most favourable (62 per cent positive), governments do not use them to produce published reports as
frequently as the slightly more critical private consultants (58 per cent positive) or even more critical academics (52 per cent supportive). Moreover, there was no trend in our survey for the researchers used more frequently by government to produce more favourable reports: those who had done only one report for government were more likely to have written a report supportive of policy than those who had written between two and five as well as those who had more than five reports (though the differences are not statistically significant).

Table 1 Positive findings in evaluation reports by authorship

<table>
<thead>
<tr>
<th>Author</th>
<th>N</th>
<th>% positive*</th>
</tr>
</thead>
<tbody>
<tr>
<td>University-based academics</td>
<td>58</td>
<td>52</td>
</tr>
<tr>
<td>Private consultants</td>
<td>58</td>
<td>58</td>
</tr>
<tr>
<td>Mix with academics</td>
<td>35</td>
<td>57</td>
</tr>
<tr>
<td>In-house researchers</td>
<td>40</td>
<td>62</td>
</tr>
<tr>
<td>Mix without academics</td>
<td>7</td>
<td>53</td>
</tr>
<tr>
<td>Unclear</td>
<td>8</td>
<td>47</td>
</tr>
<tr>
<td>Total</td>
<td>206</td>
<td>54</td>
</tr>
</tbody>
</table>

*Coded as "positive" or "very positive"
Source: coded reports (see text)

Does government use the research specification to produce positive reports?

In the pursuit of political objectives, a government department might be expected to seek to make sure that the results of the policy evaluation reflect well on their policies and their work. Governments might try and skew the research to produce favourable reports by defining carefully the research questions to be pursued and the methods used to pursue them. One does not have to look very far in the free comments and interviews to learn that specification of the research questions and methods was a common area in which the political clashed with evidence-based objectives. One respondent in a free comment argued that specifying the research methods, along with the selection of the topic for evaluation, was key in government's attempt to produce politically acceptable results:

More often however the real place where research is politically managed is in the selection of topics/areas to be researched and then in the detailed specification. It is there that they control the kinds of questions that are to be asked. This gives plenty of opportunity to avoid difficult results.

Another pointed out:

I have worked on two projects from two different government departments and the problems were very similar. In both cases, we were doing an experiment and the level of interference was high in the design of the experiment, such that in the end the answer for one of the studies was not worth having as we were not allowed to compare the government supported intervention against a realistic control group. We had to fight continually to maintain the integrity of the research design.
Government needs to be hands off when it funds research projects.
A third suggested
I was shocked at the level of interference of civil servants at certain points in the progress of the research. Specifically, they intervened at the sampling stage, changing entirely the case-study sample which we had selected to be as representative as possible of the variables involved in the overall scheme. The final sample chosen related partly to the extent of development of some agencies but it also included agencies which seemed most likely to operationalise the key features of the policy initiative. The consequence of this on the one hand was that slow starting/developing agencies which might highlight a range of contextual problems were not included. On the other hand it meant that those most sympathetic to government values and most closely aligned with the thinking behind the policy were more likely to be represented.

One interviewee offered another specific example "The sample of students chosen to do the scheme was not random but involved targeting various groups that teachers believed would have benefited most from the scheme". Another was very explicit about the kind of direction received: "We were told by the [department] to evaluate how cost effective the project [was]... [W]e were pretty much told at the beginning the purpose of the report was to show that the programme was cost effective, i.e. we had to try to show it in a good light"

However, write-in answers cannot tell us how often such departmental management of research questions and methods might be found. The survey responses give us perhaps a more direct way of evaluating the mix of evidence-based and political logics involved in government-funded policy research they offered in the case of commissioning. When we asked how "the initial research questions you were to address in the project developed", 45 per cent answered that "the government organisation(s) had a clear idea of the precise questions the research should examine", 27 per cent suggested that "the government organisation(s) had a broad idea of what the research should examine and left the definition of the precise research questions to me/us" and 26 per cent that the "precise research questions were developed jointly by me/us and the government organisation(s)" (2 per cent did not know, N=190).

Thus, nearly half the time, government gave clear instructions over the research questions. However, that government had a clear idea of the research questions is not the same as saying that it sought to impose a political logic on the research. The evidence is only slightly mixed on whether this is likely to have been the case. Of the reports that were based on questions devised by researchers alone (i.e. the questions were "left to us", N=43), only 23 per cent were supportive of policy compared with 50 per cent of the reports (N=80) where governments set the research questions. This suggests that research is less than half as likely to be supportive of government where academics rather than government sets the research questions -- indicating the possibility of substantial constraint on research findings through political objectives seeking to shape the specification of the research. The finding that where government and researchers worked together to devise the research questions (N=44) 57 per cent of the reports were supportive does not offer neat support for the idea that the degree of government
involvement shapes how kind the research will be to government policies. Rather it suggests that academics tend to be more critical of the basic idea behind the policies only when left entirely to their own devices in designing research questions. When government was involved in developing them, whether alone or in conjunction with researchers, their reports were less likely to be critical. This in turn suggests that political goals are frequently introduced in this stage of research commissioning.

Budgetary constraints can also serve to support a political objectives in commissioned research. Salisbury et al (2010) show how the constraints set by research design features specified by commissioning departments -- the budgets, the timelines as well as the specification of the methods to be used -- can prevent the generation of clear evidence-based judgments of how well or badly a policy is working. Instead, for example, of the research design allowing some form of "control group/experimental group" comparison that might be expected to determine with some precision the effects of a programme, they show how service provider comments and general indicators of satisfaction are mandated -- these generally tend to give more flattering accounts of the impact of policies than harder evaluative techniques such as Randomised Control Trials. We could not examine this important source of bias in the commissioning behaviour of government directly, but our evidence suggests it is important. When we asked our respondents whether they agreed with the statement "Keeping costs down plays too large a role when government organisations are deciding what type of research they want", 47 per cent agreed (including 16 per cent who "strongly agreed") while 23 per cent disagreed (including 2 per cent who "strongly disagreed").

The evidence suggests significant effort is frequently exerted by those commissioning research to come up with a research design that avoids embarrassing or uncomfortable results and that this attempt seems to work to a significant degree. We cannot conclude from the evidence that academics are pushed around by those commissioning them, or muzzled to produce politically favourable results as such efforts do not invariably work. As discussed above, our survey shows that even research projects where the questions were entirely set by government were as likely to be critical of policy as not. Moreover, whether such efforts reflect a clear intention to skew the research or result from a range of constraints operating on those involved in commissioning research is hard to determine. As one respondent who complained especially vehemently about the political constraints introduced in the research design stage speculated, how far the narrowing of research questions really reflected "a conscious political act or a failure of imagination or a more subtle restriction of thought I do not know" and added "it is of course still up to us as researchers whether we think enough room has been given for us to make an honest and useful contribution without compromise our own integrity. That is a decision that is also made subject to some strong institutional pressures balanced by the need to be maintain a scholarly reputation with one’s academic peers".

**Political research management?**
When research is commissioned by the department, an official or officials from the department is/are usually given responsibility for managing the research -- making sure that it is progressing, dealing with any questions that arise, handling data provision and
access to other government organisations among other things. A handful of free comments and passages in our interviews indicated the potential for the imposition of political objectives at this stage. In a free comment, for instance, one respondent said that the experience of working for government "is not one I would wish to repeat. Civil servants kept a very close eye on the research and the research process and there was a real risk that academic freedom would be compromised as a result". An interviewee complained "The liaison with the [departmental] team that was managing the programme … was difficult. The evaluation was micro-managed despite a lack of research knowledge in the [departmental] project management team".

However, while examples of this (possibly political) micro-management by officials can be found, our evidence does not suggest it is particularly frequently experienced by academic researchers. Certainly officials managing research tend, according to our academic respondents, to show significant interest in the work of their researchers; 60 per cent agreed that "officials in government organisations responsible for monitoring research generally maintain a high level of interest in it" while only 16 per cent disagreed (24 per cent neither agreed nor disagreed). A substantial contributor to perceptions of lack of interest might be the common view that "civil servants move from one post to another so often that there is too little continuity in its management of research projects" with which 59 per cent agreed and 20 per cent disagreed (21 per cent neither disagreed nor agreed, see also Salisbury et al 2010: 112). Respondents who reported a lack of interest were significantly more likely (70 per cent) than others (51 per cent) to agree that civil servants moved around too much.

However, interest in research cannot be equated with micro-management, still less that of a political type. When asked specific questions about official involvement in their most recent evaluation report, respondents did not present a picture of officialdom as a channel for political steering of their research during its progress. 17 per cent of respondents felt there was "too much contact" in the course of the project \(^3\), 4 per cent that there was too little but the large majority, 73 per cent, felt the amount of contact was "just right". Moreover the fact that the amount of contact did not appear to be substantially related to how positive the report produced turned out to be (40 per cent of the reports of those with "too much" contact were supportive compared with 48 per cent of those with "just right" amounts of contact, a statistically insignificant result) suggests that routine supervision of research does not appear to be accompanied by successful pressure to produce politically acceptable results. We will come back to this question later since, as we shall see, we have reason to believe that those officials with the more routine contacts with researchers might be less likely to press political over evidence-based objectives on researchers.

\(^3\) The question we asked was Thinking of all the contact you had with officials in the commissioning organisation over the entire project, face-to-face as well as through phone, mail, email and other electronic forms, would you say that there was Too much contact Too little contact About the right amount of contact
Political authorship of reports?
Perhaps the stage at which one would expect to find the most pressure from the ministerial clients to ensure a report is supportive, or at least less critical, would come in the report-writing stage. It is conventional for researchers to discuss a draft of a report with the client before submitting it for publication. At this stage we might expect the bureaucrats supervising the research to seek to change or scale down critical content within it. When we asked whether the government sought to make substantial changes to the draft report, 46 per cent said they were not asked to make any changes, 33 per cent were asked to make "one or two" changes and 19 per cent more than two substantial changes. The government proposes changes that affect the interpretation of findings or the weight given to them in over half the reports that the academics in our survey had recently conducted.

The free responses in the survey as well as the interview material gives ample evidence of those sponsoring the research seeking to shape the way the results are reported. If one were to judge solely on the basis of frequency of mentions of such attempted influence, the attempt to shape how the results are framed would be the single most important stage of the transaction between departments and researchers at which one would find an attempt to sustain political goals in commissioned research. A few quotes will give a flavour of the comments we received:

There were some requests for wording changes to the report so that it would be in keeping with the political agendas

There was a huge pressure to spin findings to fit their agenda

There was a lot of dialogue back and forth at the end between us and the DoH before it was published to ensure they did not look bad. They wanted certain wording changed so that it was most beneficial from a PR and marketing point of view; and they wanted certain things emphasised and certain things omitted

The individual who had demanded (despite our advice) [that we provide a particular set of figures because he felt it would make the programme look better] then wanted this analysis to be removed from both the Executive Summary and the main report - he argued that this would not put the programme in the 'best-light'. As I had been asked to provide this [analysis] four days before the final report (and had already been pulling 120 hour weeks), I refused. However, the whole [department's] project management team continued to send emails and make phone calls (approximately four or five a day) demanding its removal. In the end, this data was removed from the Executive Summary ...

---

4 The wording of the question was "When you were finalising your report, did the commissioning organisation(s) ask you to make any substantial changes to your findings? By ""substantial"" is meant any changes that affect the interpretation of your findings or the weight that you might give them"
And the people working with the departments will come and change the wording of your results without explanation; the secretary/assistant of the department would edit and remove bits and pieces…. “the minister does want this or that”….

Moreover, our survey data suggests that when asked to make some changes to the final report the academics tend either to oblige or meet their sponsors half way. When we asked what their reaction to the requests were, of those asked to make changes (N=89) only 5 per cent made no changes and 20 per cent simply made the changes asked for, 74 per cent reached a compromise with the government sponsors (1 per cent did not know).

There is thus plenty of evidence that government asks its researchers to make changes. This makes it somewhat surprising that the relationship between being asked to make changes and how supportive of the government policy the report was are by no means clear, though the small number of cases involved means that they are not statistically significant and can only be suggestive. Those (N=29) asked to make most changes (i.e. more than two substantial changes) produced a lower percentage (29 per cent) of supportive reports than those asked to make no changes at all (46 per cent). Those asked to make one or two changes produced (with 52 per cent) the highest proportion of supportive reports. While being asked to make substantial changes generally is associated with reports more critical of government, the fact that they remain critical suggests the changes did not change critical drafts into supportive final reports. And the fact that 46 per cent of draft reports approved by the department without any substantial changes were critical of policy suggests that government is tolerant of significant amounts of criticism without seeking to change anything.

Moreover the reactions to the requests by the sponsoring departments to make changes does not support the view that the attempts to subordinate the evidence-based goals of the research to political goals are particularly successful in muzzling evidence-based criticism of policies. While only 4 respondents made no changes in response to departmental requests (producing two reports that were supportive and two that were not), the 14 respondents who made the changes requested produced fewer supportive reports (14 per cent) than the 59 who compromised (53 per cent). Making changes in response to requests from the department does not appear to be associated with the probability that the report will be critical of policy.

These survey findings are somewhat puzzling and at first appearance entirely at odds with the evidence provided by the free comments and interviews of powerful government editing and sanitising research findings. The free comments threw up many examples of attempts to change the way the research report was produced. Most of them, however, also went on to suggest, broadly, that they gave in on the cosmetic changes and held firm on the substantial. As one respondent put it: “There were some requests for wording changes to the report so that it would be in keeping with the political agendas - if this went contrary to the data, we did not make the changes. When suggested wording was not contrary to the data, we did make changes”. Another pointed out:

- There was a lot of dialogue back and forth at the end between us and the
[Department] before it was published to ensure they did not look bad. They wanted certain wording changed so that it was most beneficial from a PR and marketing point of view; and they wanted certain things emphasised and certain things omitted. However they could not change the actual findings themselves. And several respondents suggested that it was in their contracts that it was ultimately up to the researchers what they wrote in the report. As one said in interview

They can’t ask you to change what you have. There are guidelines preventing them. They can’t tell you not to write something. The most they can do is tell you to change the wording, they could tell you to express something a little less strongly. [Interviewer: Does that happen a lot?] Not that often, it’s normally quite justified when it does happen. At the end of the day it is in the interests of the research group, the department isn’t obliged to publish anything.

Of course, it is possibly to be expected that academics will have this heroic self-image -- as conscientious scholars resisting pressure to spin results politically in the name of personal and professional integrity -- and seek to project it in this way. As one questionnaire respondent put it, "ultimately if the findings are not favourable some departments and the government puts pressure on to mediate findings. On the whole [private] consultancies rather than academic researchers seem happy to do this and make findings bland or toned down". The only support for this position we can offer is its consistency with the findings already discussed; that considerable pressure is frequently placed on researchers and this does not appear to be sufficiently strong to turn critical drafts into supportive reports. However, since these findings are themselves based on the perceptions and self-reported behaviour of academics we must point out that it is no more than a plausible hypothesis; a proper evaluation of the hypothesis would require a systematic textual analysis of different drafts of a report and the circumstances in which they were produced. Computer Weekly's (2006) report of the National Audit Office's scrutiny of a big health IT project highlights that it is possible to show how reports can change in the process of "clearance", though also highlights that it is easier to show evidence of a change in tone than a change in substance.

The official report is, of course, not always the only or even main route for disseminating research findings. Governments use research findings in press releases and speeches, and here it is much harder for academics to shape the interpretations of their results; these can be written without academic involvement or approval and by the time they are delivered it is too late to do much about it. Several interviews and free comments complained of the way the findings were used outside the formal report. For example, one interviewee recalled

Cameron [the Prime Minister] gave a speech [in 2012] about [a particular type of intervention]. Research had shown that this was not effective, but in the speech he mentioned only the good results, even though I had said that I did not want to make the results publicly known before the peer review. Cameron did not lie, but he only said the good things, and this is what they usually do: they use the results for news manipulation. However, they have never stopped anything from being published.

A survey respondent offered a recent experience in free comment:
There was a huge pressure to spin findings to fit their agenda (which we mainly resisted but there were compromises). Their press release following the publication of our report was highly selective in the evidence it drew upon and contained inaccuracies. By the time we pointed this out and they changed it, news outlets had already widely reported the inaccuracies. Altogether a very jading experience.

Another argued

Our experience was that government allowed our research report to stand pretty much as we had written it. However, the press release accompanying the research was written by government and substantially overclaimed what we had found. The release was not misleading, as such, it was just strong 'PR spin'. Our team decided not to engage directly with the media and this was the correct decision as we could otherwise have been forced to either disagree with our funders or to defend claims we did not make.

A very experienced researcher wrote "One area your questionnaire doesn't ask about is dissemination and particularly press releases, where we have had some problems with Govt departments wanting to edit what we send out". Though we did not explore this directly in the questionnaire, it seems that the dissemination of the research results through the press and media it is likely that the government is most able to change the tone of the research to fit political objectives and academics least able to do much about it.

**Forms of "government" influence**

We have been talking about "government" as a whole as a conduit for asserting political objectives in commissioned research. Several respondents in interviews and free comments made distinctions between different audiences for the research within government. There are at least four distinct groups within the sponsoring department that potentially have different relationships with researchers and different interests in pursuing political or evidence-based objectives. Not all four groups will be involved in every research project, but they are nevertheless worth setting out. First there are the officials who are responsible for research, possibly because they are researchers themselves. Where mentioned, these seem to have the best relationships with researchers and appear most likely to share a belief in the importance of evidence-based objectives in research. As one suggested

A key distinction in my experience is between commissioners of research and their policy counterparts. It's the latter who are often the trickier to handle, whilst the former sometimes even see themselves as protecting research integrity against the demands of the policy people. This was certainly my experience of doing work … in a politically contentious area

The second group, as this quote suggests, are the "policy people" -- officials with the task of looking after policy within the department, whether to amend, defend, expand or contract it (see Page and Jenkins 2005). Another survey respondent makes a similar point about the political zeal of the policy officials: "the research manager places a lot of emphasis on research integrity, whereas the policy teams may have their own ideological or policy motives".

The third group is the ministerial political leadership, already discussed in this paper. Our
survey evidence suggests unsurprisingly that they are highly likely, in the view of our respondents, to downvalue evidence-based objectives of research as only 4 per cent of respondents (N=182) agreed with the proposition that "Ministers are prepared to act on evidence provided by social research even when it runs counter to their views" (71 per cent disagreed and 25 per cent neither agreed nor disagreed). This is consistent with one written comment that was also reflected in others: "my area is a sensitive one, and civil servants generally are interested in research results whereas politicians have a keen eye on how the public will respond (especially The Daily Mail)". One researcher described his perception of the civil servants being on the side of the researchers against the minister: "often our results are critical of policy interventions but I find civil servants open and accepting of the findings even while they are having to think how to square the results with ministerial expectations".

A fourth group is that provided by the professionals and service providers in the programmes being evaluated. These can be brought into contact with researchers more or less directly. A particularly direct incorporation of such officials in commissioned research is through their inclusion as "stakeholders" or "experts" in steering groups. A less routinised direct method is through inviting them on an ad hoc basis to discussions about the progress of the research, discussions of the findings or draft reports. They may also be involved indirectly through means their views being solicited draft reports, provisional findings of methodological choices. Several respondents and interviewees mentioned the role of service providers as a source of constraint on their research findings. One argued

We met regularly with the Head of Research in the [Department] and also occasionally with their policymaking colleagues. One difficulty with these meetings was that they insisted representatives of the [organisation running programme being evaluated] attended. This made it quite difficult to discuss the report openly because these peoples’ livelihoods depended on the scheme. My part of the report was critical of the [programme] and I thought it inappropriate for the [department] to invite these people along. I felt it hindered honest and open discussion.

Steering group involvement was particularly awkward for one respondent

The steering groups often set up to 'steer the research' usually have their own agendas and are usually upset when for example, if the research is evaluating a policy that steering group members may have some role in, and the findings appear to challenge the policy being implemented. This can make for difficult steering group meetings where the research team is often forced to justify their credibility and the rigour of the study, and also the findings which they invariably question (if they do not confirm their own findings) when they read the draft reports. Finally, in writing up the research findings you sometimes feel that you have no say as so much is questioned and/or you are asked to rephrase particular sentences so as to put things in a more positive light.
Does government ignore bad news?
Although a central question in the analysis of government commissioned research, the degree to which government acts on evaluation reports -- amending, developing or even abandoning altogether policies on the basis of research evidence -- says less about the importance of political objectives in the enterprise than we might think. If government were only interested in research that made it look good, then we would expect government not to act on any "feelgood" advice that the commissioning process produced since much of the time such advice would simply be saying that everything was basically working well (unless it suggested ways it could work better). If government ignored critical research-based advice this would also be consistent with our expectation about government only being interested in "feelgood" conclusions. Lack of direct influence of evaluation research on policy practice is, as has already been discussed, overdetermined, and cannot be taken as particularly strong evidence about the intentions of those in government commissioning, managing and using it.

Nevertheless, our respondents were asked about the impact of their research on subsequent policy -- only 129 respondents offered a clear answer to this question (63 indicated that it was "too early to say" or did not know and 15 did not answer). Of the 129 answering, 10 per cent claimed it had a "big impact", 40 per cent a "moderate impact", 28 per cent "little impact" and 22 per cent "no impact". Around half respondents replying claimed some impact, the other half little or none. This appears, on the face of it, to suggest a greater impact for research than one might expect on the basis of the research utilization literature that usually finds it difficult to identify any research projects that have significant influence on the subsequent development of policy. However, since the question posed did not specify the type of influence (whether, say, on the policy, its implementation or on general debates surrounding the policy), one should not place too much emphasis on it as an indicator of the influence of research on policy.

However, the degree to which perceptions of influence might be related to how critical of government policy could cast light on the importance of political objectives as government might be expected to act on friendly policy advice and be hostile to other forms (see Salisbury et al 2010: 8). In fact those writing reports that were broadly supportive of the objectives behind the policy were no more likely to claim they influenced policy than those who were not (each were 50 per cent likely to claim influence); those who were critical of the way policy is implemented were somewhat less likely to claim influence (45 per cent) than those who were supportive (62 per cent -- a difference just about significant at the .05 level) suggesting that remaining sympathetic to those implementing policy might more likely to help your report gain influence than being critical of them. But even this finding is hardly a striking example of the political uses of commissioned research.

Conclusions
We have so far examined the possibility for political objectives being introduced into the research process at discrete stages of research. The reliance on direct manipulation in research commissioning was suggested in one free comment:

The quality of research management within [the Department] has deteriorated
progressively, and the skills and professional knowledge of government research managers has fallen as their impulse to intervene and control has grown. This is partly why they increasingly favour commissioning consultancies to mount their research, rather than research centres in universities. However, there are other ways the commissioners of research can shape the conduct, conclusion and presentation of research that can best be described as through anticipated reactions. Such a mechanism is notoriously difficult to analyse, however several of the free comments and interview responses hinted at it:

I'd characterise officials' desire to shape the outcome of 'independent evaluations' as per persistent subtlety, rather via hard dictat or coercion.

We don’t want our research to be obsolete, so we shift our focus to helping make the policy better rather than suggesting that it be abandoned.

Overall positive experiences. They [policy makers] are open to criticism so long as it is not completely contrary to their overarching goals and principles. Does this moderate the extent to which we criticise? Not really but it does affect the way we word or present our criticisms

No pressure was put on us to change the findings. For the sake of future initiatives, we deliberately wrote it up as 'what could be learned from what went wrong' rather than a catalogue of what went wrong. Thus we need to recognise that our examination of direct means of seeking to assert political objectives may well underestimate the significance of political objectives. By the same token, however, we must also entertain the possibility that politicians' and civil servants' propensity to assert political values in commissioning and managing research might itself be reduced by the anticipated reactions of the researchers to attempts to shape their conclusions or how their research is to be interpreted.

Our study also was conducted at a time in which we caught many projects that were commissioned under a Labour government and did not finish until a new Conservative-Liberal Democrat Coilaition was in office. The free comments included many statements about how things had deteriorated under the Coalition. For instance one wrote

My responses are based upon research commissioned by the last government but, for the most recent project, the report was submitted to the current government. This meant that the final report was 'buried' along with others in historical research that was effectively disowned on the Departmental website. Circulation of these later reports to local partners was also impeded by the closure of regional Government Offices. Since the change of government, in [the Department I often work for] at least, it appears that there is no longer any interest in commissioning research or developing an evidence base.

We do not know precisely how many were started and finished under the Coalition as we only asked about the year in which the research was completed. However, we know that 56 were completed under Labour before 2010, 61 were completed in 2010 and were
likely to have been commissioned by Labour, we expect that a large proportion of the 85 completed in 2011 and 2012 were also commissioned under Labour although completed under the Coalition. However, there were no noticeable differences in the responses for any of the questions based on these three time periods that approached statistical significance. For example, on the variables where we might have expected strongest differences they were slight: those asked to make "more than one or two changes" to draft reports increased from 14 per cent for those finishing before 2010 to 21 per cent for those finishing in 2010 and 19 per cent for those finishing after 2010; those claiming influence for their reports was 48 per cent, 53 per cent and 55 per cent respectively for the pre-2010, 2010 and post-2010 completion groups.

The production of evidence for policy making is by its nature a political process. But just as the political process of problem solving cannot be based exclusively or even predominantly on "evidence", neither can the production of evidence be based on "science" alone. As applied research, government evaluations of policies are likely to be able to serve a range of practical purposes. One of these purposes might be to generate material for an open-minded evaluation of whether what government is doing is working and how it might do it better; other purposes might include political advocacy or public relations. The fact that the research process involves several distinct groups within government -- including research officials, policy officials, officials who run the programmes being examined as well as politicians -- each likely to have different hopes and expectations from an evaluation, means that within government itself there is likely to be found a jumbled range of range of evidence-based and political objectives (not to speak of the career-based objectives discussed in the introduction). Garbage-can theory (Cohen, March, and Olsen 1972) suggests that the objectives likely to gain ascendancy is subject to change throughout the life of the process of commissioning research as different groups within government become involved in it -- one might expect less direct ministerial involvement, for instance, in the development of the research methodology than in developing the press release accompanying the publication of a report. To find "political" objectives in the conduct or presentation of research does not preclude also finding a more open-minded search for evidence as well.

We found evidence of government seeking to pursue political objectives at all stages in the policy process. In the selection of topics our admittedly rather indirect evidence does not suggest that government only commissions research that is likely to produce "safe" or uncritical reports. In developing the research, government involvement in specifying what precisely is investigated and how appears to affect the likelihood of the research producing conclusions critical of policy, but even so, half of the respondents claimed their reports were critical even where the government effectively set the research questions on its own. Significant efforts appear to be made to get researchers to change their reports at the drafting stage. However the pressure does not appear sufficiently strong to be able to muzzle a critical first draft. Researchers often have little direct involvement in drawing up the press releases and statements that government uses to present its interpretation of the results, and these are, unsurprisingly, likely to be where the evidence of using science for political objectives is strongest.
Academics' ability to resist pressure to steer the results might appear to be strong, but we have to bear in mind that this assessment is based on academics' own perceptions. While we have no special reason to think their reported perceptions of government influence and other features associated with commissioned research are erroneous or misleading, one cannot ignore the fact that academic reputations in part rest on the ability to maintain, or at least appear to maintain, scientific standards against government interference. Academics not only have an interest in resisting political pressure, but also an interest in appearing to be able to resist it even if they cannot or do not. Yet the evidence suggests that the whole process of government commissioning and managing research by academics leaves sufficient space for academics to produce critical reports and that governments can tolerate and manage such reports without losing face or credibility.

Given that the benefits of evaluations are hard to account for in an instrumental way (i.e. that evaluations lead to improvements in the policies they evaluate), then it is hard to complain about the political expectations of politicians and the officials that work for them: without such expectations it is arguable that few evaluations would ever be conducted. If there is a good chance that any report we pick up has been written under some pressure to produce favourable or non-embarrassing results -- some direct, some indirect and some self-imposed by researchers anticipating the reactions of those paying them -- what is the value of any evaluation picked at random, aside from any political uses it might have? The value of even the most rigorous evaluation usually lies in the more indirect it may enhance our understanding of how policies and interventions work and inform debates and deliberations about policy change. The value of the research becomes a matter of how far it can make this indirect contribution. In debates about how policy works, even research that might have been designed to find a pre-established position can have value if it contains a serious attempt to weigh up this position against alternatives and/or if it provides evidence or data that can be used by others to do this. Only if it does neither could one consider it junk.

References


Levitas, R (2012) "There may be ‘trouble’ ahead: what we know about those 120,000 ‘troubled’ families" Poverty and Social Exclusion in the UK Policy series: Working paper No.3 (Bristol: University of Bristol)

Lindblom, CE and Cohen, DK (1979) Usable knowledge: Social science and social problem solving


Talbot, C (2012) " Louise Casey and ‘Listening to Troubled Families’: an (almost) worthless piece of ‘research’ leading to dangerous policy prescriptions" (Blog entry) http://whitehallwatch.org/2012/07/18/louise-casey-and-listening-to-troubled-families-an-


