

Policy Practitioners' Accounts of Evidence-Based Policy Making: The Case of Universal Credit

MARK MONAGHAN* AND JO INGOLD**

**Social Sciences, Loughborough University, Brockington Building, Loughborough LE11 3TU, United Kingdom*

email: m.monaghan@lboro.ac.uk

***Business School, University of Leeds, Maurice Keyworth Building, Woodhouse Lane, Leeds, West Yorkshire LS2 9JT, United Kingdom*

email: j.ingold@leeds.ac.uk

Abstract

This paper draws on insider accounts from UK Department for Work and Pensions (DWP) officials to analyse the relationship between evidence and policy making at a time of rapid policy development relating to Universal Credit (UC). The paper argues, firstly, that evidence selection within the DWP was constrained by the overarching austerity paradigm, which constituted a *Zeitgeist* and had a significant bearing on the evidence selection and translation process, sharpening the focus of policy officials and analysts on the primacy of quantitative evidence when advising Ministers. Secondly, while methodological preferences (or an 'evidence hierarchy') impacted on evidence selection, this was not as significant as practitioners' perceived capabilities to handle and develop evidence for policy. These capabilities were linked to departmental structures and constrained by political feasibility. Together, these dimensions constituted a significant filtration mechanism determining the kinds of evidence that were selected for policy development and those omitted, particularly in relation to UC. The paper contributes to debates about the contemporary role of evidence in policymaking and the potential of the relationship between future evidence production and use.

Introduction

In recent times, much commentary has been made on the role (or deficit) of evidence in the formulation of policy. Discussions tend to specifically focus on the kinds of evidence that make the grade in terms of their policy utility and those that tend to be ignored in the policy maelstrom. The steadily accruing literature concerning evidence-based policy making (EBPM) are indicative of these developments. Central to the EBPM literature, particularly that which takes a critical look at the evidence and policy connection, is a fundamental concern *over what* evidence, ideas and knowledge are integrated into the policymaking process.

With some notable exceptions, (for example, Cameron *et al.*, 2011; Stevens, 2011; Maybin, 2015) what is often missing from accounts of *how* evidence is or is not used by policymakers is a closer consideration of the attitudes, actions and aptitude of practitioners at the coal-face of policy development towards evidence utilisation. These are personnel that Merton (1945) some time ago referred to as ‘intellectual bureaucrats’; officials on the inside of the policy process who exercise advisory and technical functions. This paper draws on research with such a group within the Department for Work and Pensions (DWP) in the UK at a time of rapid policy development in relation to the ambitious rollout of Universal Credit (UC). UC was the Conservative-Liberal Democrat Coalition Government’s flagship reform to simplify the benefits and tax credit systems by introducing the UK’s first single working-age benefit.

Our aim was to garner an insider account of how the policy that was developed in this context shaped the evidence that was used and to understand why this was so. The paper argues, firstly, that evidence selection within the DWP was constrained by the overarching austerity paradigm, which constituted a *Zeitgeist* that had a significant bearing on the evidence selection and translation process, sharpening the focus of policy officials and analysts on the primacy of quantitative, economic data and evidence when advising Ministers. Secondly, unlike other studies that suggest that policy makers prefer certain kinds of evidence on a hierarchical basis on the grounds of perceived scientific rigour, we found that DWP officials favoured certain kinds of evidence based on their own individual capabilities to compile evidence and perceptions about what Ministers wanted in the context of departmental team structures. We elaborate on these points below. Together, these dimensions constituted significant filtration mechanisms (Stevens, 2011) that determined the kinds of evidence that were selected for policy development and those that missed out. Our findings suggest a more complex process than that depicted by a simplified ‘evidence hierarchy’ as described in extant literature and the paper raises important issues about the contemporary role of evidence in policymaking and the potential of the relationship between evidence production and use in the future.

The paper proceeds as follows. The next section considers the broad relationship between evidence and policy. This is followed by presentation of the methods employed in the study. The findings are then presented in three thematic, but in many ways overlapping, sections: (i) whether there is an evidence hierarchy within the DWP (whereby certain kinds of evidence are favoured due to their perceived methodological sophistication); (ii) reflections on the capability of respondents to handle and develop evidence for policy, and, in turn, whether and how Ministers understand this; and (iii) political feasibility, relating to the kinds of evidence deemed pertinent by Ministers and officials in the policy area at a specific point in time (*Zeitgeist*). Based on the findings, our attention turns to the future of policy development and the evidence-base for Universal Credit.

We strike a note of caution, highlighting how the policy apparatus of the DWP is at present potentially inhospitable to recent findings that have considered the human and social costs of welfare reform. We do, however, offer hints of optimism about the direction of policy travel. Finally, some concluding remarks are made about how reforms to the policy making process, including the development of criteria for evidence selection and use, may be beneficial in the future.

The impact of the policy apparatus on evidence

Emerging from debates in evidence-based medicine, evidence-based policy making became a dominant frame for policy formulation after the election of the New Labour government in 1997. This was part of a broader ‘modernising agenda’ for government (Cabinet Office, 1999; Mulgan, 2005). Cameron and colleagues note (2011: 430) that at the heart of this development was a ‘belief that policy making needed to move away from the subjective use of anecdote to a more objective, dispassionate review of available evidence’. In recent times a pragmatic turn (Pearson, 2010) has emerged in discussions around EBPM. The sense of optimism that accompanied the first UK New Labour government’s assertions that it would be guided more by ‘evidence’ of ‘what works’ than ‘dogma’ has dissipated (Nutley *et al.*, 2007; Monaghan, 2011), not least as the complexity of the evidence and policy connection became more apparent (Monaghan, 2010). The UK Conservative-Liberal Democrat Coalition government (2010–2015) made little pretence of basing policies upon evidence. This was to some extent justified by the prevalence of the narrative of ‘austerity’ as a game-changing context for the development of policies (see Taylor-Gooby, 2012; 2016).

Although austerity still shapes the social policy landscape, EBPM has retained some currency within Government. A solid portfolio of guidance from the UK central government on the process of ‘evidence informed policy’ (Bochel and Duncan, 2007) has been produced. HM Treasury’s *Magenta Book* (HM Treasury, 2011) provides guidance to policy makers and analysts for the design and management of evaluations of government projects, policies, programmes and the delivery (2011: 7). The updated version ‘shifts emphasis away from the “analyst’s manual” of the previous edition, to a broader guidance document aimed at both analysts and policy makers at all levels of government, both central and local’ (2011: 5). The *Magenta Book* is a companion to HM Treasury’s *Green Book* (HM Treasury, 2003), which constitutes binding guidance for departments and government agencies with regard to the economic appraisal of the development of policies and programmes. Appraisal and evaluation (and in particular cost-benefit and cost-effectiveness analyses – p. 4) is a key stage of the ‘ROAMEF’ cycle (Rationale, Objectives, Appraisal, Monitoring, Evaluation, Feedback) within Westminster policymaking and intended to represent the UK policymaking cycle, although in recent years it has become less central to this process.

Despite these attempts to provide guidance for policy makers and government analysts (economists, social researchers, statisticians and operational researchers) regarding the harnessing and utilisation of evidence, the publications have arguably added to, rather than alleviated, this void on the grounds that the guidance is, at best, incomplete and highly abstract. For instance, the Cabinet Office (2008) framework for international policy comparisons omits the crucial step of ‘how to transfer/translate policies’ between contexts and at no stage does the Magenta Book consider how policymakers can assimilate knowledge from such evaluations. Additionally, very little attention is given to the basic reporting of evidence to Ministers except to draw, very briefly, upon Vaughan and Buss’ (1998) tips on reporting social science research to busy policy makers.

A key limitation of the EBPM literature concerns, therefore, the relative paucity of data that explains *how* evidence is used by those charged with the task of using it. In one of the most enduring studies of the evidence and policy relationship, Weiss (1979) highlights how the relationship between evidence and policy can be seen in instrumental, in conceptual or in symbolic ways. The ‘instrumental’ version, which assumes that scientific knowledge has a direct, linear impact on policy making, has now largely been discredited. The ‘conceptual’ understanding suggests that knowledge offers opportunities for policy makers to change the way they conceive of policy issues and so in this version evidence filters into decisions over time in an ‘enlightenment’ capacity. The ‘symbolic’ version is where evidence is harnessed by policy makers to justify pre-existing policy stances rather than acting as a guide for action.

This account, although useful, remains abstract. There is, however, an emerging body of research that has attempted to consider some of the main mechanisms by which research is filtered into and out of the policy process. A common thread of the more recent accounts of evidence utilisation is that evidence selection in policy making is an unequal competition as the framing of the policy in question often determines the kinds of evidence that are deemed permissible. Liebling *et al.* (2017: 11) note how:

‘research has shown how the politics of knowledge production confound a linear conception of the movement of research into policy and practice . . . Governments seek out research to support a view already taken, rather than critically engaging with a wider field of knowledge (some of which may contradict or confound political perception) . . .’

Similarly, in their comparative review of diverging tobacco control policies in the UK and Japan, Cairney and Yamazaki (2017) note that:

‘key actors do not simply respond to new information; they use it as a resource to further their aims, to frame policy problems in ways that will generate policy makers’ attention, and inform technically and politically feasible solutions that policy makers will have the motive and opportunity to select. This remains true even if the evidence seems unequivocal. . . .’

For Cairney and Yamazaki, the sensitive nature of the policy itself impacts on the kinds of evidence that are deemed viable for policy development in the UK and Japan. Stevens (2007a; 2011) has suggested that to understand how evidence is translated into policy one must consider a number of filtration processes that render evidence selection an unequal competition. Stevens (2007a) notes how, in contentious areas of policy development, there are often significant socially structural, as well as politically tactical, barriers to evidence use. Citing the examples of immigration and drug use, Stevens notes how the direction of policy in the UK in both these areas over recent times has been towards restricting these activities. This is because those who stand to benefit from 'liberalising' the legislation (drug users and would-be immigrants or asylum seekers) are often among the least powerful members of society. They are also among the groups frequently scapegoated and stigmatized for social problems, particularly crime, making political reform less likely (see Ford and Lymeropoulou (2016) for a discussion of immigration and crime and Seddon (2005) and Stevens (2007b) for the links between drugs and crime).

Elsewhere research (Ingold and Monaghan, 2016), has suggested that several, inter-related, factors broadly shape the process of translating evidence into policy. Firstly, that the substantive nature of policy issue or problem determines the kind and extent of evidence used. Secondly, agenda setting and policy framing influence the search for permissible evidence (Head, 2008). Thirdly, specific filtration processes (Stevens, 2007a; 2011) are always at play in the relationship between evidence and policy. Here, the kinds of evidence that are deemed useful to policy makers tend to survive, whereas critical evidence falls by the wayside. Fourthly, following lessons from institutionalism, where political institutional arrangements greatly affect policy processes and outcomes (Parkhurst, 2016: 9) the apparatus for policy design and implementation can have a direct bearing on the kinds of evidence used and omitted. Fifthly, policy personnel in the form of analysts or policy officials play a decisive role in relation to each aspect.

Despite this emerging literature, there are, as indicated, few studies that have directly analysed the routes by which evidence finds its way into policy from the point of view of those who facilitate or are responsible for this. One consequence is that evidence hierarchies are often used as an explanatory tool to explain the preponderance of the use of certain kinds of evidence in policy over others (Nutley *et al.*, 2013). Our research sought to move beyond such accounts, in order to (i) shed light on the dimensions that impact on evidence utilisation; (ii) better understand how evidence is employed in the policy process; and (iii) analyse views of evidence utilisation from a user perspective. Although we would argue that the five factors identified above are inter-related, for current purposes in this paper we investigate to what extent (i) policy itself is a determining feature of evidence selection and utilisation, specifically within (dominant) policy frames; and (ii) policy shapes the kinds of evidence that are deemed permissible and silences

other, potentially valuable, contributions. We were particularly interested in how the organisational features of the DWP at the time of our research mediated the framing of the evidence and how this was influenced by the broader policy landscape and times (*Zeitgeist*). Before turning to the findings from our research, we give below a brief account of the (very politicised) policy context in which our research was conducted.

Research context

The Department for Work and Pensions (DWP) is the UK's largest administrative department and is responsible for pensions, social security benefits and child maintenance policy. It oversees employment policy and delivers employment services both through contracted provision and through Jobcentre Plus. Three workshops were conducted with 75 DWP policymakers and analysts in three of the department's key locations around the UK. Workshop participants were officials and analysts (social researchers, statisticians and economists), ranging from middle-ranking to senior officials and from relatively new entrants to officials of long service. Participation in the workshops was voluntary through self-selection but, from their everyday roles, all participants had an interest in how evidence and policy interconnect. The DWP was a compelling location for this case study given that between 2004 and 2010 it oversaw an unprecedented growth in commissioned evaluations (Legrand, 2012: 330). However, Legrand's assertion that DWP's growth in research outputs demonstrated an 'extraordinary sea change in their use of research evidence in policy development and evaluation' assumes that this evidence was actually used in policymaking. During the research period, the DWP was engaged in a wider agenda of experimentation with new approaches across government, including 'horizon scanning', which looks across a spectrum of evidence sources to consider how they could shape future policy developments. These initiatives represented a continuation of the machinery of government changes led by the UK Cabinet Office but were also aligned with political imperatives under austerity, such as large-scale expenditure cuts to central government departments and drives for greater efficiency. Consequently, the workshops took place during a time of important organisational change and significant policy change for the DWP.

The most important aspect of the latter was the policy development of, and preparation for, the introduction of Universal Credit (UC). The central aim of UC is to 'radically restructure' the existing benefits and tax systems by merging six benefits and tax credits into one single working-age benefit, in order to improve incentives to work (DWP, 2010). UC was introduced in April 2013 in pathfinder areas (pathfinder rather than pilot given that definite national rollout of the policy was already decided), with final completion across the UK expected by 2017. These expectations have not been met. The reform

has attracted much media coverage and criticism, with repeated claims that the project had floundered and had been ‘scaled back’ (Wintour, 2013) and more recent revelations about the reality of income loss and the winners and losers of the policy (Savage, 2018; BBC, 2016). There have been concerns around the financial difficulties for households resulting from the change from weekly to monthly payments, financial support for childcare and problems with online accessibility. Other concerns include UC’s implications for women, particularly the individualisation of benefit conditionality without individualised payment incentives for second earners, who are more likely to be women (Gingerbread, 2013; Millar and Bennett, 2016;). The National Audit Office has published critical assessments of the implementation of UC (2013; 2014) and the government’s own Infrastructure and Projects Authority (IPA) recommended a ‘firebreak’ in implementation in early January 2018 in order to consider whether the rollout plan remained appropriate. The Work and Pensions Select Committee (2018) recently praised the DWP for bringing the project ‘back from the brink of complete failure’ but highlighted concerns regarding ongoing challenges (2018: 14) relating to this ambitious policy reform.

UC is the result of repeated political attempts to address the difficult task of simplifying the UK’s complex social security and taxation systems. Early in the first governing period of New Labour, incremental change was favoured over more large-scale overhaul, leading to the introduction of Tax Credits and changes to the benefit rules concerning spouses/partners of benefit claimants (Joint Claims for Jobseekers Allowance for couples and the New Deal for Partners). Further changes to social security entitlement followed in the Welfare Reform Act 2009, increasing conditionality for a wider range of individuals. Central actors in the next phase of reform were the Centre for Social Justice (CSJ) and Iain Duncan Smith who was, at the time of our research, the Secretary of State for Work and Pensions. The CSJ first published its UC blueprint for the report *Dynamic Benefits* (CSJ, 2009), focusing on (dis-)incentives to work in the existing social security system, in particular high marginal tax rates. Government plans for UC were then set out in the Green Paper *21st Century Welfare* and, finally, in the White Paper *Universal Credit: Welfare that works* and enshrined in the Coalition Government’s Welfare Reform Act 2012. These documents drew on a relatively limited range of evidence and, notably, not upon findings from DWP-commissioned evaluations, including those relating to the New Deal for Partners or New Deal for Lone Parents in terms of the real-world complexities of moving into work and changes to circumstances beyond technocratic or rational economic dimensions.

At the time of the workshops, DWP was still finalising aspects of UC and discussion of this important policy change and its affiliated evidence-base pervaded each of the workshops. Workshop discussions related to the overall policy development of UC, its implementation with HMRC (responsible for the management and collection of tax and payment of tax credits) and with

local authorities (who administrate Housing Benefit, another payment merged into UC), as well as in relation to DWP-commissioned evaluations of the policy. Drawing on research by one of this paper's authors in three countries in collaboration with the DWP (Ingold and Etherington, 2013), discussion in the first workshop largely related to consideration of UC in relation to couples and their decision-making in relation to work and benefits and 'In Work Conditionality', now referred to as 'In Work Progression' (IWP). IWP is a critical dimension of UC relating to its central aim to incentivise *any* work¹ by extending benefit conditionality to those in work but still in receipt of benefits. IWP is without precedent internationally, with a very limited evidence base (the exception being DWP's Employment Retention and Advancement demonstration project focused on improving job retention and progression (Hendra *et al.*, 2011)). Consequently, there has been a series of DWP-commissioned IWP pilots to garner what could be termed 'policy-based evidence'. Additionally, the first DWP-commissioned UC evaluation report was published in 2011, focused on public perceptions of welfare reform and UC (Rotik and Perry, 2011). This was followed by the first pathfinder evaluation (2013) and estimations of UC's employment impacts based on a limited claimant group in 2 per cent of areas; qualitative and cognitive research including 'laboratory' experiments utilising random assignment (DWP, 2017) (underscoring the dominant role of 'nudge' and the Behavioural Insights Team in policy post-2010); and evaluations of mixed method 'test and learn approaches' in relation to families. Within this policy context, the workshops aimed to tease out policy makers' and analysts' understandings of the evidence and policy connection.

Our research was co-produced in consultation with the DWP and resulted in three workshops, the first workshop comprising 10 participants, the second around 25 participants and the final workshop almost 40 participants. The workshops ranged in length from one to two hours. Although there were internal variants, there were constants in the subject matter discussed. In each scenario, we presented conceptual models based on the EBPM and policy transfer/translation literatures as vignettes to stimulate discussion. The key research question posed was: 'How is evidence used, or translated, in the context of contemporary policy making in the DWP?' All workshops were transcribed verbatim and data were analysed thematically. The findings are presented thematically in the following three sub-sections: (i) Evidence hierarchy?; (ii) Capability; and (iii) Political feasibility.

Findings

Evidence hierarchy?

Reflecting the emergence of evidence-based policy from evidence-based medicine (Solesbury, 2000), Pawson (2006: 49) highlights how there has been a

‘fairly seamless’ adoption from health into other policy areas of a ‘gold standard’ of the randomised controlled trial (RCT) (preferably with concealed allocation). Research that is seen to be ‘scientific’ and able to establish causality with strong external validity tends towards the summit of the hierarchy, while languishing near the foot are studies based on small samples, for example qualitative case studies, as well as the professional knowledge and opinion of practitioners and stakeholders.

In social policy, RCT studies are less common than in health or medicine, for example, but the primacy of quantitative evidence exists over other, more qualitative, forms. Ultimately, the hierarchy is based on judgements over methodology and the perceived ‘scientific credentials’ of the method by the policy maker. Although in our research we found evidence of methodological preference similar to a hierarchy of evidence, it was not immediately clear that this was about the superiority of one method over another. Instead, it seemed to be based more on the skill set of officials and what they deemed was practically and politically feasible. In one of our workshops officials discussed their experiences of using survey data to guide policy, stating that:

‘Survey data has limitations...Samples can be skewed. Response rates can be low. It’s a messy story!’

The following participant also underlined how, because the relationship between evidence and policy is ‘messy’ (a point to which we return subsequently), there are particular limitations within certain policy areas, such as employment policy:

‘... random control trials are at the top of the hierarchy but you run a RCT in a recession it will give you a different answer than if there is growth. So in my experience sometimes there is a clear direction but sometimes the evidence points this way and sometimes points that way. There are lots of examples, child poverty, Universal Credit where we do draw on evidence and data but it isn’t a straightforward story’

Within UK employment policy, the use of RCTs has been particularly contested, largely due to the ethics involved in exposing one group to an intervention at the expense of others (this is not the case in other countries, for example Denmark). In the UK there have only been two examples of RCTs in employment policy: the Joint Retention and Rehabilitation Pilot (Purdon *et al.*, 2006) focused on providing interventions to facilitate the return to work of employees on long-term sickness absence and, more recently, the Employment Retention and Advancement demonstration programme (Hendra *et al.*, 2011), mentioned above.

The quote below expands on the discussion concerning the roles of varying types of evidence in the policy process in respect of administrative (management information) data held by the DWP and the complementary role of surveys. It resembles the distinction highlighted by Maybin (2015) who, drawing on the work of Tenbesel (2006), states that policy work involves a variety of knowledges:

episteme – rational analytical knowledge; *phronesis* – the deliberation over values; and *techne* – the knowledge about how to ‘do’ policy work:

‘administrative sources are more complete and you might want a survey on what views are at the time. We do have some facts. There are lots of examples of policies that draw on the data but it is complex. Admin data lacks social data. We also have to work with Ministers’ own views. There are some facts, such as the decline in pension scheme enrolment but we don’t know how effective policies to address this are. Evidence based policy is the ideal but it is messy and complicated. That is the challenge. We know what the problems are but we can’t for example force people to take out pensions. We can set up automatic enrolments, but people may choose to opt out. We do try, I like to talk about evidence based policy but the reality is that it’s messy.’

Similar to Smith and Joyce’s (2012) work on understanding attempts to develop evidence-based policy in public health, our participants frequently suggested that evidence was ‘messy’ and ‘complex’ and pointed to a role for both quantitative and qualitative evidence. However, as the following quote demonstrates, they did not necessarily perceive a primacy of quantitative over qualitative evidence, which Smith and Joyce (2012) note in their study.

‘qualitative evidence is more important in some aspects than quantitative.’

This perhaps suggests that there is the potential for a broad range of evidence based on different methodologies to find its way into the policy process. However, the next section discusses some of the limitations on this.

Capability

In our research the issue of capability related to, on the one hand, participants’ individual reflections on their own capabilities in terms of both handling and developing evidence and, on the other, their perceptions of Ministers’ capabilities to understand research. These were, in turn, impacted by Departmental structures. A consistent finding from our study was that the officials we spoke to were buoyed by their capacity to be able to harness evidence. At times, however, this seemed at odds with academic accounts (e.g. Mulgan, 2005) of what evidence is and can achieve in the policy process:

‘I think in the academic literature there’s generally a consensus and you can work that out pretty quickly. Any analyst in the department, even a junior one could be given a job like that, straight out of university you could do something like that. I did something on parenting and family structure – in a couple of days you can work out what the key findings are on that’

This view of capability is commendable but it is also revealing of views of evidence and its relationship to policy from the inside. In some ways, this contradicts the earlier views of the inherently complex nature of evidence and its associated methodologies. The previous quotation was by no means representative of all the analysts’ views, in that it points towards a rational, linear policy process where evidence exists and can directly be inserted into policy discussion, circumstances

permitting. Other appreciations from our research (as we shall see) and with other literature tend to highlight the complexity of the evidence and policy connection (Sanderson, 2009; Smith and Joyce, 2012). For example, in their review of the state of the art of meta-analysis, Rosenthal and Di Matteo (2001: 60) demonstrate how, in nearly every field of research, new findings ‘daily overthrow’ older ones, meaning that findings are ‘often confusing and conflicting about central issues of theory and practice’. In effect, the more something is researched, the less agreement there is and the concept of ‘consensus’ in academic literature is seldom borne out.

There is, of course, the argument that one would expect to see some confidence and optimism about evidence from those invested in trying to design and implement policy. Nonetheless, evidence is clearly contingent on a conducive context for its use:

‘ours is more a policy/analytical division so the evidence is maybe closer to me in terms of how it works in reality. So, there does seem to be more of an ongoing conversation around influencing things.’

Here the capability of evidence use is contingent on officials’ role within the organisation – analysts will be ‘closer’ to the evidence than their policy colleagues for example and so are committed to its use. The commitment to evidence is, in turn, often contingent on stability within teams and divisions. One recurring finding was the importance of collective institutional memory on the delivery of evidence for policy.

‘It’s different depending on where in DWP you work. There are different structures. In some cases analysts sit alongside the policy teams and work together and in other cases it’s more fragmented.’

There is, however, the perennial issue of resourcing, which came to the fore in one of our focus groups when discussing how analysts decide what evidence to use.

‘I think some of it’s a time issue because realistically there’s not that many of us and there’s not a huge amount of time always so it could just be things that we know we’ve done internally, or from talking to other people who might know things. It would be lovely to trawl the evidence base every time you have a question but we can’t do it. Actually, with the In Work Conditionality this is one of the better examples because there’s been a bit more time to think about things. But often it tends to be what’s the organisational memory and things like that.’

Although, as is clear from the previous section, participants did not necessarily subscribe to the primacy of quantitative data over qualitative (even if they felt that Ministers did), in the workshops they also expressed concern regarding their lack of confidence about mixed methods. This suggests that participants were aware of the potentially rich picture that could be provided by a broad range of evidence but had concerns over their capability to organise this evidence for Ministers.

A significant weight of evidence to shift perceptions can be difficult to manage and has a clear knock-on effect for the feasibility of officials to harness evidence for policy. This accords with Rose's (1991: 24) assertion that in policy transfer there are two key standards against which programmes and policies should be judged: technical feasibility ('Is the programme practical?') and political feasibility ('Is it desirable?'). These two aspects were prominent in our research in relation to officials' perceptions of capability (individual and departmental), but were also shaped by a third dimension: the context (or political 'Zeitgeist') in which policies were being designed and implemented.

Political feasibility

'you have the evidence and you have the policy and sometimes the policy doesn't reflect the evidence, whereas how I see it evidence is just one of the influences on policy. There's a whole load of other stuff ... 100 per cent evidence-based isn't right because of what I was saying about the triangle before – feasibility, political will, there's legal frameworks, there's all sorts of things. Some policies will be much more evidence-based than others ...'

The political 'Zeitgeist' for the research was the then-Coalition government's overarching programme of austerity and the purported need to drive down public spending on social security (predominantly concentrated in the area of working-age benefits). Taylor-Gooby (2016) illustrates how this era has seen the welfare state move from an instrument of social cohesion to one of social division. Taylor-Gooby (2016) shows how various reforms since 2010, specifically the increase in value-added tax (VAT) from 17.5 per cent to 20 per cent and the lowering of income tax 'shifted the tax burden downward' (2016: 718), universal VAT being particularly regressive. This context underlined the significance of the role of the economy and economic data (Ingold and Monaghan, 2016), but also calls into question claims that evidence-based policy is non-dogmatic:

'Some policies will be much more evidence-based than others will be so the benefit cap, you could have done that as an evidence based policy and looked at what's the right level analytically or you could have done it on ... it's ideologically based, that you shouldn't get more on benefits than you do as a medium-earning family.'

The significance of the policy Zeitgeist added a further level of complexity in that it shaped what kinds of evidence were thought to be feasible or, in other words, palatable to Ministers. A key issue of feasibility related to what analysts perceived policy officials and Ministers wanted and would accept as evidence. Here an element of gatekeeping could occur:

'There's also capacity of policymakers – if you have policymakers who ... don't quite grasp it, it's quite hard to get across to them the complexity. And especially people often want nice clear findings as well so complexity is quite hard – I suppose complex findings are quite hard to use when you're trying to design a policy it's quite hard if you're told "This is all really complex."

This quotation reflects a longstanding issue in the evidence and policy literature: the extent to which there are two separate communities and how compatible they are (e.g. Caplan, 1979; Head, 2010). This is an enduring feature of discussions about evidence-based policy because clearly, in this context, policy makers' expectations shape the kinds of work that analysts do.

A further key finding from the research was that temporal aspects had a specific impact on the feasibility of analyst teams to compile evidence for policy and, in turn, this impacted on their appreciations of how they handled the evidence-base. Here 'timing' can relate to both the time-frames in which policy actors are obliged to work, but also 'the times' in which the policy is being implemented. The latter is better thought of as the policy *Zeitgeist*, which influences the framing of the policy problem and shapes the substance of the response. The following quotation neatly encapsulates the interplay of these factors, drawing on the example of conditionality in the benefit system:

'Evidence is very much at the heart of it. Evidence, operational feasibility and the political. But that depends on other stuff like the economy... policy generally evolves, it moves along, you can't generally see it happening, but incrementally every day it's changing a bit. I think someone said earlier that the big policy changes aren't evidence-based. Maybe the idea isn't, but the implementation is really focused on evidence. The evidence might be that we want to do X, but then the evidence comes in and we say 'How do we achieve what we want to achieve?' I feel fairly positive about how evidence is used where I work.'

This point about which kinds of evidence drive policy reveals how it is often not the technical kind of number crunching (episteme), but the evidence which maps onto politics and electioneering and 'playing to the gallery' (phronesis). As indicated, this theme directly influenced perceptions of feasibility and capability but what was most noticeable was a trade-off between capability and feasibility. This is also demonstrated by the quote below:

'Universal Credit... my own impression is that we're only addressing half of the policy. There's no general policy of creating enough jobs in the right places for all the ways we are pushing people into the workplace. So UC is about making people better off in work and there's the increasing pension age which is dumping pensioners so the way it looks is that we'll be driving down wages because there aren't enough jobs for people to do. Again that's the economic situation.'

This quote is telling in that it provides both a critique of the policy direction in the context of austerity and post-recession, as well highlighting the balancing of 'winners and losers' of policies when connecting evidence and policy.

Conclusions

Unlike other studies that suggest that policy makers have a preference for certain kinds of evidence on the grounds of its perceived scientific rigour, in our research we found that DWP analysts and policy officials felt that ministers were more persuaded by the findings of large-scale, national social surveys than by other

kinds of evidence. Officials were also more comfortable handling and processing this kind of evidence in terms of their own individual technical capabilities. The context of departmental team structures also had a bearing on this capability and, furthermore, this was also aligned with what officials saw as being politically feasible in relation to Ministers' preferences, which added a layer of complexity into the selection and use of evidence. In this way, the policy apparatus was a key determinant of evidence use. Our findings are semi-consistent with what has been referred to as the hierarchy of evidence in policy making and thus mirror the findings from other studies of evidence and policy (Smith and Joyce, 2012).

Our research also suggested that evidence selection within the DWP was constrained by the overarching austerity paradigm, which constituted a *Zeitgeist* and a further significant evidence filtration mechanism. The study fieldwork was conducted in the infancy of Universal Credit (UC), which, from its outset, has generated much media coverage and criticism. Our findings suggest that, despite officials' acknowledgements of the limitations, there was a reliance on numbers and what could be 'objectively' measured. This sharpened the focus of policy officials and analysts on the primacy of quantitative evidence when advising Ministers, in line with the capabilities outlined above. As such, austerity as a dominant policy frame was a key determining feature of the (non)selection and (non)utilisation of evidence by analysts and policy officials in the DWP, manifesting in a preference for economic, quantitative data.

With no clear guidance on how evidence is to be used in policy, and with perceptions of capability and feasibility seemingly driving evidence use, there is a danger that vital, emerging evidence will continue to be obscured or omitted in this process. For example, it is unclear how research relating to a Minimum Income Standard required for an acceptable living standard covering material needs (Davis *et al.*, 2016) is accounted for within the complex emerging picture of recognisable winners and losers under UC (Hirsch and Hartfree, 2013). In addition, evidence of the human (rather than purely economic) cost of austerity such as the widespread need for foodbanks (Garthwaite, 2016) or the deleterious impact of sanctioning on benefit claimants (Patrick, 2014; 2017) have not yet filtered into the policy process. The same could be said of other research that challenges the current austerity paradigm² (Miller and Bennett, 2017). Yet this research provides clear lessons for policy.

An Institute for Government (2012) report acknowledged that the DWP 'was traditionally strong on analysis' (2012: 19) but that the supply of evidence was a secondary issue to that of its demand (2013: 16–17). To develop effective policies, policy makers and analysts need to look beyond the constraints of outputs and outcomes in the form of numbers in relation to a specific policy area. While the analysts themselves recognised complexity, this tended to be in relation to social research methods. In this way, other kinds of complexity are obscured, including the tensions of overlapping and often contradictory policy areas,

including for different groups of recipients under a policy with a wide reach such as UC.

We do not intend to merely resort to calling for more resources in the form of expenditure, but it is essential that government policy makers and analysts have the necessary resources and skills to better equip them with the capability to interpret, present and translate a broad range of evidence into policy, including a wealth of evidence that already exists. This relates to how policy and analyst teams are structured in order to facilitate better utilisation of evidence within departments in central and local governments (as well as government agencies). It also involves the development of criteria for good evidence *use*, in addition to existing guidance about evidence *selection*. Finally, it also requires better integration of a wider range of epistemic communities into the policymaking process through more porous boundaries between government and academia.

The complex and potentially path-breaking context brought to the fore by recent political developments and the emerging evidence-base raises many questions. Are analysts in DWP still committed to developing robust evidence for policy and, if so, how do they achieve this? Will we continue to observe a preference for numbers among analysts and policy officials in DWP or will other forms of evidence be required, and used? These can only be answered by the practitioners themselves. The changing political environment (including Brexit and devolution) may give voice to new forms of evidence but, unless more attention is placed on how evidence is translated into policy, this will continue to be mediated by the technical capacity of practitioners, as well as more arbitrary notions of what is considered feasible within existing public policy structures.

Acknowledgements

We would like to thank the anonymous reviewers for their helpful comments on the article. The research was carried out with a small grant from the Higher Education Innovation Fund.

Notes

- 1 Once in work, UC recipients who are below an earnings threshold would be required to increase their working hours to the equivalent of a 35 hour week, for example through a combination of additional employment, higher hourly wages, increased hours, or an alternative job.
- 2 <http://www.welfareconditionality.ac.uk/about-our-research/>

References

- BBC (2016), 'Universal Credit leaves working families worse off, says IFS' <http://www.bbc.co.uk/news/business-35475350>
- Bochel, H. M. and Duncan, S. (Eds.). (2007), *Making policy in theory and practice*. Bristol: Policy Press.

- Cabinet Office (1999), *Modernising government*, CM4310, London: The Stationary Office
- Cabinet Office (2008), *Beyond the horizon: International comparisons in policymaking*. London: Centre for Management and Policy Studies.
- Cairney, P. and Yamazaki, M. (2017), A Comparison of Tobacco Policy in the UK and Japan: If the Scientific Evidence is Identical, Why is There a Major Difference in Policy?. *Journal of Comparative Policy Analysis: Research and Practice*, 1–16.
- Cameron, A., Salisbury, C., Lart, R., Stewart, K., Peckham, S., Calnan, M., Purdy, S. and Thorp, H. (2011), Policy makers' perceptions on the use of evidence from evaluations. *Evidence & Policy: a journal of research, debate and practice*, 7(4), 429–447.
- Caplan, N. (1979), The two-communities theory and knowledge utilization. *American behavioral scientist*, 22(3), 459–470.
- Centre for Social Justice (2009), *Dynamic benefits: Towards welfare that works*. London: Centre for Social Justice.
- Davis, A., Hill, K., Hirsch, D. and Padley, M. (2016), A minimum income standard for the UK in 2016, York: Joseph Rowntree Foundation.
- Department for Work and Pensions (2010), *Universal Credit: Welfare that works Cm7957*. London: DWP.
- Department for Work and Pensions (2017), *Understanding how Universal Credit influences employment behaviour: findings from qualitative and experimental research with claimants*, available online at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/643952/understanding-how-universal-credit-influences-employment-behaviour.pdf
- Ford, R. and Lymperopolou, K. (2016), Immigration: How attitudes in the UK compare with Europe, Findings from the British Social Attitudes Survey, available online at http://www.bsa.natcen.ac.uk/media/39148/bsa34_immigration_final.pdf
- Garthwaite, K. (2016), *Hunger pains: Life inside foodbank Britain*. Policy Press.
- Gingerbread and the Children's Society (2013), *Single Parents and Universal Credit: singled out?* London: Gingerbread.
- Head, B. W. (2008), Three lenses of evidence-based policy. *Australian Journal of Public Administration*, 67(1), 1–11.
- Head, B. (2010), Evidence-based policy: principles and requirements. *Strengthening evidence-based policy in the Australian Federation*, 1, 13–26.
- Hendra, R., Ray, K., Vegeris, S., Hevenstone, D. and Hudson, M. (2011), Employment Retention and Advancement (ERA) demonstration: Delivery, take-up, and outcomes of in-work training support for lone parents. Research Report 727. Sheffield: DWP.
- Hirsch, D. and Hartfree, Y. (2013), *Universal Credit: making work pay*, York: Joseph Rowntree Foundation
- HM Treasury (2003), *The Green Book: appraisal and evaluation in central government*, London: HM Treasury
- HM Treasury (2011), *The Magenta Book: guidance notes for policy evaluation and analysis. London: HM Treasury (Magenta Book Background Papers)*.
- Ingold, J. and Etherington, D. (2013), Work, welfare and gender inequalities: an analysis of activation strategies for partnered women in the UK, Australia and Denmark. *Work, employment and society*, 27(4), 621–638.
- Ingold, J. and Monaghan, M. (2016), Evidence translation: an exploration of policy makers' use of evidence. *Policy & Politics*, 44(2), 171–190.
- Institute for Government (2012), *Evidence and evaluation in policymaking a problem of supply or demand?* London, Institute for Government.
- Legrand, T. (2012), Overseas and over here: policy transfer and evidence-based policy-making. *Policy studies*, 33(4), 329–348.
- Liebling, A., Maruna, S. and McAra, L. (2017), Introduction: The New Vision, in Liebling, A., Maruna, S. and McAra, L eds. *The Oxford Handbook of Criminology*, 6th Edition, Oxford: Oxford University Press

- Maybin, J. (2015), Policy analysis and policy know-how: A case study of officials in England's department of health. *Journal of Comparative Policy Analysis: Research and Practice*, 17(3), 286–304.
- Merton, R. K. (1945), Role of the intellectual in public bureaucracy. *Social Forces*, 405–415.
- Millar, J. and Bennett, F. (2016), 'Universal Credit: Assumptions, contradictions and virtual reality. *Social Policy and Society*, 16(2), 169–182.
- Monaghan, M. (2010), The complexity of evidence: Reflections on research utilisation in a heavily politicised policy area. *Social Policy and Society*, 9(1), 1–12.
- Monaghan, M. (2011), *Evidence versus politics: Exploiting research in UK drug policy making?* Bristol: Policy Press.
- Mulgan, G. (2005), Government, knowledge and the business of policy making: the potential and limits of evidence-based policy. *Evidence & Policy: A Journal of Research, Debate and Practice*, 1(2), 215–226.
- Nutley, S. M., Walter, I. and Davies, H. T. (2007), *Using evidence: How research can inform public services*. Policy press.
- Nutley, S., Powell, A. and Davies, H. (2013), *What counts as good evidence? provocation paper for the alliance for useful evidence*. London: Alliance for Useful Evidence.
- Parkhurst, J. (2016), *The politics of evidence: from evidence-based policy to the good governance of evidence*, London: Routledge
- Patrick, R. (2014), Working on welfare: Findings from a qualitative longitudinal study into the lived experiences of welfare reform in the UK. *Journal of Social Policy*, 43(4), 705–725
- Patrick, R. (2017), *For whose benefit?: The everyday realities of welfare reform*. Bristol: Policy Press.
- Pawson, R. (2006), *Evidence-based policy: a realist perspective*. Sage.
- Pearson, M. (2010), 'What do we know? What should we do?' Melding research validity and rhetoric in the analysis of policy making. *Evidence and Policy*, 6(1), 77–90.
- Purdon, S., Stratford, N., Taylor, R., Natarajan, L., Bell, S. and Wittenburg, D. (2006), *Impacts of the Job Impacts of the Job Retention and Retention and Rehabilitation Pilot Rehabilitation Pilot Rehabilitation Pilot*. Research Report 342. Leeds: DWP Corporate Document Services.
- Rose, R. (1991), What is lesson-drawing? *Journal of Public Policy*, 11(1):3–30
- Rosenthal, R. and Di Matteo, M. R. (2001), Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual review of psychology*, 52(1), 59–82.
- Rotik, M. and Perry, L. (2011), Perceptions of welfare reform and universal credit, *Department for Work and Pensions Research Report no 778*, available online at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/214563/rrep778.pdf
- Sanderson, I. (2009), Intelligent policy making for a complex world: pragmatism, evidence and learning. *Political Studies*, 57(4), 699–719.
- Savage, M. (2018), Ministers told to expect backlash as millions lose out from Universal Credit, available online at <https://www.theguardian.com/society/2018/apr/07/universal-credit-backlash-millions-lose-out>
- Seddon, T. (2005), Drugs, crime and social exclusion: social context and social theory in British drugs–crime research. *British Journal of Criminology*, 46(4), 680–703.
- Smith, K. E. and Joyce, K. E. (2012), Capturing complex realities: understanding efforts to achieve evidence-based policy and practice in public health. *Evidence & Policy: A Journal of Research, Debate and Practice*, 8(1), 57–78.
- Stevens, A. (2007a), Survival of the ideas that fit: an evolutionary analogy for the use of evidence in policy. *Social Policy and Society*, 6(1), 25–35.
- Stevens, A. (2007b), When two dark figures collide: Evidence and discourse on drug-related crime. *Critical Social Policy*, 27(1), 77–99.
- Stevens, A. (2011), Telling policy stories: an ethnographic study of the use of evidence in policy-making in the UK. *Journal of Social Policy*, 40(2), 237–255.

- Taylor-Gooby, P. (2012), Root and branch restructuring to achieve major cuts: The social policy programme of the 2010 UK coalition government. *Social Policy & Administration*, 46(1), 61–82.
- Taylor-Gooby, P. (2016), The Divisive Welfare State. *Social Policy & Administration*, 50(6), 712–733.
- Vaughan, R. J. and Buss, T. F. (1998), *Communicating social science research to policymakers* (Vol. 48). Sage Publications, Inc.
- Weiss, C. H. (1979), The many meanings of research utilization. *Public administration review*, 39(5), 426–431.
- Wintour, P. (2013), Auditors challenge government over cost of universal credit IT problems, London: *The Guardian*, available online at <https://www.theguardian.com/politics/2013/dec/10/auditors-challenge-cost-universal-credit-it-problems> (accessed 20 April 2018).