

## Understanding Evaluation in the UK

**Abstract** This chapter provides a genealogical tracing of the discourse of evaluation and governing. Evaluation is frequently presented as characteristic of the ascendance of evidence-based policy-making (EBPM). However, through the lens of youth work in the UK and a discussion of the shifting attitudes of policy-making elites towards evaluation research, the chapter disrupts this discourse. As an alternative, it argues that evaluation's emergence as part of governing has been facilitated by the growth of New Public Management and developments in information communication science. The chapter reviews these discourses, showing that themes within NPM and ICS have been absorbed by evaluation and the evaluation–governing relationship.

**Keywords** UK policy-making · EBP · Youth work · New public management · Information communication technology

### INTRODUCTION

The central argument of this book is that evaluation is representative of a specific style of governing which involves disciplining and controlling subjects through knowledge production processes. Before engaging in this discussion, I first want to revisit how evaluation has been operationalised in the UK context (using youth work as the policy context) and indicate what evaluations and the attitudes of policymakers to evaluation

can tell us about the role of evaluation in governing. Additionally, adopting a Foucauldian genealogical approach, I want to draw attention to the discourses which have impacted and are manifest in discussions relating to evaluation and the understandings and position of evaluation in policy. This will provide an insight into how evaluation is constructed in policy—the role attributed to it and the reason for its prominence—beyond the narrative of EBPM. This narrative asserts that evaluation is prized by governing elites as it provides necessary evidence to inform decision-making and implies that prominence of evaluation is due to the respect governing elites have for EBPM. Challenging this implication, I want to argue that evaluation's primary function is to manage policy subjects and communicate information about their activities. Its prominence in policymaking was facilitated by developments in public management—specifically the adoption of New Public Management (NPM)—and expansions in information and communication technologies (ICT).

### THE OPERATIONALISATION OF EVALUATION IN YOUTH WORK

In his celebrated review of the relationship between auditing and governing *The Audit Society*, Power (1997) argues that although discussions on the need for government programmes to be audited took place in the late nineteenth century, auditing has only become a feature of policymaking cycles in the United Kingdom relatively recently. While the first statute on auditing in governing was introduced under William Gladstone's premiership in the 1890s, auditing has only really become visible and active since the 1960s. Moreover, its position in the terrain of governing was peripheral until the establishment of the National Audit Office and Audit Commission under the Thatcher governments of the 1980s. As with auditing, evaluation has only become a distinguishable part of the policymaking cycle relatively recently. While at this point there are numerous professional organisations and journals dedicated to evaluation, as well as government-produced frameworks for where evaluation should sit in the policy-cycle and demands that evaluation should be part of all policymaking and implementation, as Pawson (2013) notes, evaluation remains something of a cottage industry.

The late introduction of evaluation to the policy sphere in the United Kingdom is reflected in youth work in England and Wales. While there had been some reviews of youth work by the Department of Education before and after the Second World War—the Albermarle Committee

produced the first distinguishable evaluation of the youth work sector in England and Wales in 1960. Established in 1959 by the Minister for Education, the Albermarle Committee were given the responsibility to appraise and record the wide range of activities that fell under the broad parameters of youth work. Since the mid-eighteenth century, organisations had emerged to support the social and moral development of young people. The Youth Service—which was only latterly recognised as a sector of the welfare state in the mid-1940s, by which point youth work organisations such as the Boys’ and Girls’ Brigades had been in operation for almost a century—was and is still very much a ‘mixed economy’, including philanthropic, state-funded and state-run organisations. Like its service-user population, youth work—both as a sector and a practice—defies a bounded definition and, by critics and supporters, has been described as closer to an ‘art’ than a format or category (Young 2002).

As the first major review of youth work, the majority of the Albermarle Report was directed to:

[Reviewing] the contribution which the Youth Service of England and Wales can make in assisting young people to play their part in the life of the community, in the light of changing social and industrial conditions and of current trends in other branches of the education service, and to advise according to what priorities best value can be obtained for the money spent. (HMSO 1960, para. 1)

As an evaluative document itself, Albermarle is very much resonant with an experimental approach to policymaking or evaluation as planning and design (similar to a ToC approach). It is not a summative appraisal of a large body of evidence—as the Report states, the Committee did not ‘[undertake] any large scale research projects in what is a very wide field’ (HMSO 1960: para. 3). That is not to say that Albermarle is not an evaluation or the integration of research with policymaking. It presents a robust appraisal of the impact of youth work to the lives and social and educational development of young people and a broad audit of staffing, training and resources in the youth work sector in England and Wales. These findings are translated into a generalised criticism of the state of the sector along with 54 recommendations for restructuring, centralised management and public financing. It also advocated the establishment of the Youth Services Development Council (YSDC) to monitor and

coordinate the delivery of a 10-year development plan to ‘bring the service up to date’ (Milson 1970: 15).

The next significant sector-level evaluation was commissioned by the YSDC in 1967. This evaluation was completed by two study groups: the Fairbairn Committee which focused on the relationship between the Youth Service and schools; and the Milson Committee which focused on the role of the Youth Service in community development. The individual reports of the Committees were amalgamated in a single evaluation report—*Youth and Community Work in the 1970s* (the Fairbairn-Milson Report). This was presented to parliament in 1970. As a piece of research, Fairbairn-Milson was much more rigorous than Albermarle. Each committee went to great efforts to collect and analyse information on the ‘state’ of the Youth Service and youth work in England and Wales. It offered both a summative assessment of gaps in service and the merits of various permutations of youth work and made recommendations on how to develop and improve the service. Fairbairn-Milson therefore seems much closer to EBPM. At the same time, like Albermarle, Fairbairn-Milson is very much a ‘planning’ document. The majority of the report is dedicated to developing and structuring effective youth work, with efficacy defined in terms of numbers of service users and impact on youth transitions to employment and active citizenship.

Albermarle and Fairbairn-Milson provide an interesting insight into the function of evaluation in policymaking and the relationship between evaluation and governing. Albermarle was commissioned as an evaluation. It had been instructed to appraise youth work, identifying the merits and state of the Youth Services, highlighting areas of inefficient and inadequate performance. However, the Albermarle Report emerged as somewhere between a distillation and a manifesto and was not prefaced by a rigorous exploration or analysis. While the Fairbairn-Milson Report was produced through a process of evidence collection and analysis much closer to EBPM, it too focused primarily on communicating a general picture of youth work and orienting the sector towards particular goals and operations. Evaluation within these is constructed as something to be used to enable more effective program design, planning and implementation rather than the systematic production of evidence on which to base policy decision-making.

The other reviews of the Youth Service between 1980 and 2000 replicate this interpretation of evaluation as part-communication of what an

evaluand is or does and part-management of an evaluand. Of the nine reviews of different elements of youth work in the United Kingdom, two reports stand out as major, wide-range evaluations of the contribution and impact of the sector—*Experience and Participation, Review Group on the Youth Service in England* (the ‘Thompson Report’—1982), and *Effective Youth Work. A Report by HM Inspectors* (1987). What is striking about these evaluations is, like Albermarle, while they position themselves as evaluations of the merits of youth work (and so fit into a robust evaluative typology), their predominant contribution to youth work policy was describing positive activities present in the Youth Services and orienting the Youth Service towards a particular mode of organisation. This is reflected in *Effective Youth Work*’s outline of the operations and organisation of the Youth Service.

‘Evaluation as description and management’ is both further reflected and entrenched in the key post-2000 evaluation of youth work—the House of Commons Education Committee (HCEC)’s 2011 *Review of youth services*. Launched in October 2010, the HCEC review was presented as an enquiry into the provision and achievements of the Youth Services. A key objective of the review was to assess how the Youth Services could be made more effective and efficient and improve both access to and use of youth services by young people. Interestingly, in terms of understanding the relationship between evaluation and governing, both the initial HCEC review and subsequent Committee responses to the government’s commentary on their recommendations presented the introduction of evaluative frameworks as a mechanism for ensuring impactful youth work provision.

Following the HCEC review’s assertion that the impact of youth services could be enhanced through evaluation and monitoring, a framework document—*A framework for positive outcomes for young people* (Young Foundation 2012)—was published. Commissioned by the Department of Education and produced by the Young Foundation, with the support of a consortium of youth work-oriented agencies, local authorities and private sector firms, this document outlines a model for conducting ‘scientifically rigorous’ evaluation of youth work’s merit which can be used to improve practice. This was a progression of the relationship between evaluation and youth work from the Albermarle, Fairbairn-Milson and Thompson periods and established evaluation as a mechanism for managing youth work. The framework focuses on social and emotional capabilities organised in a seven-point cluster and sets

out ‘a matrix of available tools for measuring these capabilities [and] outlines a step by step approach in practice’ (Young Foundation 2012: 4–5). Central to this evaluation matrix is the identification of desired outcomes and outputs by programme managers at the beginning of the intervention. The rationale for this is threefold. According to the Young Foundation report, a priori identification of outcomes and outputs will:

- *Clarify what the programme is trying to achieve (content) and how (process)*
- *Establish where the programme is working well and where further improvements are needed*
- *Close the loop with feedback on progress against business needs* (Young Foundation 2012: 24)

The framing of ‘scientifically rigorous’ evaluation as involving the pre-identification of outcomes and outputs is further entrenched by the Cabinet Office’s flagship programme for improving evaluation in youth work: the Centre for Youth Impact (CYI). Launched in November 2014 and proposed as a support agency for youth work professionals looking to develop their research skills, the CYI’s objective was to ‘provide overarching support for all impact measurement initiatives that are relevant to the youth sector’ (Cabinet Office 2014). Importantly for this, despite claiming that the methodological frameworks they suggest to professionals will be ‘bespoke’, the first stage to conducting a good evaluation according to Project Oracle, one of the three key organisations leading the CYI, involves identifying desired goals and outcomes. This is made explicit on Project Oracle’s own advisory website where they explain evaluation as requiring the clear identification of outcomes (Project Oracle website). Project Oracle argues that the ‘validity’ of an evaluation can be assessed by comparing the evaluation against what they call the Standards of Evidence (Project Oracle website). An evaluation that meets the minimum standards of scientific rigour, they argue:

has provided a coherent and plausible description of the logic that lies behind it. This includes a description of the project activities, intended outcomes and aim, how these are connected and what assumptions are being made. (Project Oracle 2014)

In terms of unpacking the relationship between evaluation and governing, the youth work sector provides interesting insight. The limited

number of sector-level evaluations also makes this relationship easier to explore. It illustrates that evaluation is used in youth work primarily to organise and direct rather than inform policy and practice through identifying productive ways of working. While evaluations of the sector have highlighted examples of good practice, youth work policy has never been produced following or as a result of the findings of an evaluation. As Jeffs (1978) notes, even before the Albermarle Report had been delivered to parliament, the shape of the Youth Service for the 1960s was being designed. This places a question mark over the accuracy of the construction of the evaluation–governing relationship as a manifestation of political decision-making based on post hoc examination of evidence (the narrative of EBPM). Whereas methodological writing on evaluation, such as that referenced in the introduction, presents evaluation in policy as intended to critique and illuminate the merits of a particular approach through rigorous scientific analysis, the reality of policy evaluation in the United Kingdom was very different with political decision-making often entirely separate from evaluation research.

### ATTITUDES TO *EVALUATION*

The treatment of evaluation by policymakers and their attitude towards evaluative processes is also at odds with the notion that contemporary governing is based on facts and evidence of merit. The capacity of UK policy evaluations to comment on the merit of particular policy approaches or suggest alternative policy arrangements—something Scriven (1981) considers a key aspect of policy evaluation research—was intentionally restricted by successive governments from the mid-1970s to the mid-1990s. As Pollitt (1993) describes:

Indeed, it sometimes seemed that the government was actually opposed to the evaluation of the impacts of its policies, especially if that evaluation was to be conducted by independent agents. Its dominant concept of evaluation appeared to be the kind of quick, narrow-focus analysis provided by a management consultant. (Pollitt 1993: 356)

After some early efforts to embed evaluation research in strategic planning at a local and national level in the early 1970s through the establishment of the Central Policy Review Staff (under the Heath government) and Programme Analysis and Review system, the capacity of evaluations to make judgements on policy was steadily constrained.

Under the Thatcher governments from 1979 to 1992, planning-focused government-coordinated policy evaluation was not embraced; both the CPRS and PAR were scrapped before the mid-1980s. The National Audit Act (1983), which led to the establishment of the National Audit Office and the Audit Commission reduced policy research to auditing and restricted the capacity for evaluation research oriented towards ‘judging and improving’ (the US approach) significantly. The NAO in particular were legally prohibited from questioning ‘the merits of policy objective’ (House of Commons Library Research Division 1992: 34).

The reticence of UK governments before 1997 to adopt evaluation as a means of improving public administration is explained by Robinson (in Ryle and Richards (eds) 1988) as symptomatic of a general unwillingness to look too closely at programmes and/or policies. In stark contrast to the US public administration’s advocacy of systematic reviews and evaluations and the narrative of professionally certified, independent evaluative practice, members of the Public Accounts Committee (the Select Committee to whom the NAO and Audit Commission reported and the *de facto* government coordinating body of policy evaluation in the UK):

[had] neither the skills nor the inclination to probe deeply into the management of public services except where there are obvious horror stories of incompetence and mismanagement. (Robinson, in Ryle and Richards (eds) 1988: 154)

That said, the use of evaluation as management but not commentary on merit or production of evidence to inform policy-design was, at face value, challenged by governments towards the end of the twentieth century. The Blair/New Labour government elected in 1997 in particular have been positioned, by themselves and UK evaluation science (Sanderson 2000; Pawson and Tilley 1997; Talbot 2010), as the standard-bearers for policy evaluation research in the United Kingdom, epitomised by the promotion of a ‘what works’ and evidence-based approach to policymaking and practice (EBPM). Unlike the Conservative governments of Thatcher and Major who were, at best, unenamoured with policy evaluation, New Labour invested heavily in coupling evaluation with governing. The quintessential statement of this was made by David Blunkett, then Secretary of State for Education and Employment, in a 2000 address to the Economic and Social Research Council where he claimed:

[S]ocial science should be at the heart of policy making. We need a revolution in relations between government and the social research community – we need social scientists to help determine what works and why, and what types of policy initiatives are likely to be most effective. (Blunkett 2000: 36 citation from Sanderson 2000)

It is important to not understate the shift in attitude towards evaluation that occurred between the end of the Major government and the commencement of the New Labour parliament. As Sullivan describes, ‘the scope and scale of the commitment to evaluation and evidence was out of proportion to the much lower-level approach to policy evaluation practised by the previous Conservative administrations’ (Sullivan 2011: 502). This shift in attitude towards evaluation has been addressed by commentators on public management and social and public policy as characteristic of the emergence of evidence-based policy and practice (EBP)—also known as EBPM—as a central doctrine of governance and practice across the domains of government—education, health, social welfare and housing—from the mid-1990s. Enabled by developments in and global networks of evidence-based medicine such as the Cochrane Collaboration, the EBPM-movement projected itself as closing the gap between research and practice (Reynolds and Trinder 2008). Central to EBPM is ‘the brave assumption that the truth will out, and that is it possible to provide dispassionate, independent and objective evidence to evaluation policy options’ (Pawson 2006: 7) and ‘the expectation [...] that policy-makers [...] will summon reviewers to have a close look at the evidence *before* the leap into policy and practice’ (ibid: 8).

EBPM promotes a circular model of policymaking (Cairney 2012, 2016) where the introduction, continuance, rejection or adaptation of policy and policy programmes is dependent on and directly preceded by systematic evaluation of:

...the extent to which the policy was successful or the policy decision was the correct one, if it was implemented correctly, and, if so, if it had the desired effect. (Cairney 2012: 33)

Under the umbrella of New Labour’s ‘pragmatic’ politics of EBPM (Clarke 2004a, b; Pawson 2006), flagship governmental programmes such as the ‘New Deals’ for employment (the New Deal for Lone Parents, the New Deal for Young People, and the New Deal for

Disabled People), the New Deal for Communities, the Health Action Zones and the Local Strategic Partnerships all had continuous, systematic evaluation of micro-level operations and outcomes built into their design. At a more general level 2004 *Green Book on Appraisal and Evaluation in Central Government* produced by the Treasury provided a model of the policy cycle—labelled the ROAMEF (Rationale, Objectives, Appraisal, Monitoring, Evaluation, Feedback)—cementing the centrality of policy evaluation to New Labour’s approach to governing. Moreover, the evaluation guidance provided by the Treasury and embedded in New Labour’s flagship programmes emphasised ‘what works’ and the need for a dynamic relationship between evaluation and governing.

From 1997, UK policy evaluation thus appeared much closer to US approaches. Micro-theories of ‘what works, for whom, under what circumstances’ (Pawson and Tilley 1997) and systems theories of the dynamics of social programmes (Sanderson 2000) were adopted ‘to explore what happens in the “black box” between programme inputs and outputs’ (Sullivan 2011: 503). In a radical change from the relationship of the 1970s and 1980s, evaluation was used formatively in the *design* of policy and evaluators were encouraged to include judgements on the merits of policy approaches and suggest potential adaptations using evidence from pilot studies (Squires and Measor 2005; Taylor and Balloch 2005). Impact case studies and cost-effectiveness, VFM evaluations, while still prominent in the government’s definition of evaluation, began to be counterweighted with formative and theory-based evaluative frameworks such as ToC, complex systems evaluation, and Realistic Evaluation (Treasury 2004, 2006; House of Commons Education Committee 2011) which focused on patterns of functionality and dysfunctionality—things working and not working—within the programme implementation and operational *process*. Evaluation was increasingly put to work as a continuous cycle of action, reflection and feedback with systematic review at its core (Pawson 2006; Weiss 1972; Young et al. 2002).

At the same time, it is equally important to not *overstate* the agentic power of evaluators under New Labour or their capacity to interject evaluative approaches outside econometric auditing into policy research. Although the evidence-based practice movement encouraged policy research to move outside the narrow parameters imposed by successive Conservative governments pre-1997 and made overtones towards action

and participant action research, the capacity for critical policy evaluation research of the kind Guba and Lincoln (1989), Greene (2009) and Mertens (1999) advocate was still very limited. There was still a great deal of scepticism among policymakers about the usefulness of evaluation research in assisting programme development and, as MacDonald (1976) emphasised 20 years earlier, policy evaluators were still directed by policymakers. The contingent nature of Realistic Evaluation and complexity evaluation in particular made them something of a hard sell for pragmatic, responsive governing elites. Additionally, the time needed to conduct a systematic review or evaluation and the orthodoxy of its pre-, during- and post-intervention structure (Pollitt 2014) was out of sync with the norms of policy change which were based on ideological arguments (Stone 2002), ‘deeper, less visible, political presuppositions’ (Fischer 2003: 14), and changed rapidly.

Furthermore, despite the extensive overtures made by the New Labour government towards rigorous, systematic policy evaluation research, no independent body with enforcement capabilities was established. Organisations such as the Audit Commission changed little from what they were in the later 1980s and early 1990s—‘influential [...] but [...] toothless’ (Jenkins and Gray 1990: 66). Recommendations from evaluations under New Labour were just that and it was up to policymakers whether to actually absorb evidence produced by evaluators into their policies. This begs the question whether there was significant change in attitude or whether the discourse of EBP was simply lip service (Jenkins and Gray 1990). This is not a new situation for evaluation to find itself in. The political agency of evaluation research (as an instrument for improving policy) has long been questioned. For example, writing on the function of evaluation of UK educational policy and practice in the 1970s, MacDonald (1976) argued that what distinguished evaluators from other social science researchers most of all was their *lack* of agency in the research process. Evaluators, in MacDonald’s view, were not able to alter (or indeed criticise) the programmes they were evaluating:

The position of the evaluator is quite distinct, and much more complex [than the academic researcher]. The enterprise he is called upon to study is neither of his choosing nor under his control. He soon discovers, if he has failed to assume it, that his script of educational issues, actions and consequences is being acted out in a socio-political street theatre which affects not just the performance, but the play itself. (MacDonald 1976: 130–131)

Leaving aside New Labour's rhetoric of EBPM and policymaking through the use of evidence from evaluations by experts, post hoc reviews of UK evaluation in the period 1997–2010 (a period characterised by intense expansion of policy evaluation's role) suggest that the shift in the *practice* and *application* of evaluation research was in reality quite limited (Sullivan 2011). This point is discussed in some depth by Pawson (2006) who notes that, despite its expansion at an industrial scale, one of the most noticeable characteristics of the New Labour approach to policymaking was 'the failure of evaluation research to feed significantly and successfully into the policy process' (Pawson 2006: 8). Similar arguments are raised by Hammersley (2013) and Pollitt (2014), the former of whom argues that the dominance of EBPM is a myth and the latter of whom characterises public management in the United Kingdom as a system of rapid, mass reform with little evidence to support changes.

Additionally, underneath the claims about challenging the 'audit culture' dominance and enlivening more multimodal, mixed methods research (particularly qualitative research and post-positivist methodologies), policymakers still preferred 'findings that could be easily translated into universal policy messages' and commissioned evaluations 'that could be sure to provide some form of statistical output' (Sullivan 2011: 506). Such a preference towards output-oriented economic findings that could feed directly into policy is reflective of global discourses of 'good evidence' in policymaking (Denzin 2009; Denzin and Giardina 2008). In the UK the *realpolitik* of evaluation (as output studies) is demonstrated by a renewed interest in performance measurement and the increasing use of target measures (Hood 2006; Bevan and Hood 2006) and the extension of the jargon of impact measurement across the public realm (Clarke 2004a, b). The result of this was that UK policy evaluation remained very close to auditing and monitoring rather than (re)designing policy through systematic research. More innovative aspects of evaluation research theory and science (participatory methods, transformative research, democratic research) were all left behind as policymakers adopted 'classical' (economic and positivist) evaluative frameworks.

Sullivan's arguments place a question mark over the extent to which the rhetoric of EBPM and theory-based, responsive evaluation research translated into practice. If her arguments regarding the preference to performance indicators and linear studies are correct then, despite the

investment, there is little difference between evaluation now and evaluation in the early 1990s which would ‘gravitate towards evaluation designs which would produce focused, practical studies which possessed contemporary policy relevance [...]’ and were ‘not designed to catch *unintended* effects of a programme’ (Roberts and Pollitt 1994: 542). This lack of difference, Sullivan explains, is due to a UK preference towards ‘concrete factual realism’ or ‘unvarnished verisimilitude’ which has translated into ‘hierarchy of methods that regard randomised controlled trials as a “gold standard”’ (Sullivan 2011: 507–508).

Consequently, despite the ambitions and opportunities presented by ‘theory-based evaluation’, UK policymakers (with the support of many evaluators) revised and refined theory-based approaches in order that they may be employed in the pursuit of ‘concrete factual realism’. (Sullivan 2011: 508)

This reorientation of evaluation models intended to be responsive and take account of unintended consequences towards positivist scientific approaches which dealt in ‘methodological fundamentalism’ (St. Pierre 2002), is also visible in the context of youth and community work evaluation. As the statements of the HCEC review indicate, despite claiming an openness to mixed methods and multimodal evidentiary forms, the HCEC prioritised what they called ‘structured’ evidence, dismissing the qualitative submissions they received as ‘anecdotes’.

Sullivan is not alone in questioning the difference between the audit cultures of the 1970s and 1980s and the expanded theory-informed evaluation science of the 1990s and 2000s in real terms. The limited application of theories of evaluation—particularly those of a more participatory and qualitative inclination—is a consistent theme across critical analyses of post-1997 policy evaluation research and the discourse of EBPM (Hammersley 2013; Taylor and Balloch 2005; Torrance 2008; Ord 2014). Across a variety of policy arenas (including education, health, criminal justice and housing), critical evaluation and critical policy studies has accused the UK administration of hollowing out evaluation research, emptying it of its orientations towards critical, action-oriented social science (Guba and Lincoln 1989) and reducing it to a ‘black box’ methodology (Pawson 2006, 2013; Chen 1994) which smacks more of ‘scientism’ (Torrance 2008) than science.

The basis of these critiques is supported by the guidance on evaluation produced by the UK government. In the 2004 Green Book, the first manual on evaluation published since *Policy Evaluation: A Guide for Managers* in 1988 (Treasury 1988), evaluation is defined as:

...a robust analysis, conducted in the same manner as an economic appraisal, and to which almost identical procedures apply. It focuses on conducting a cost benefit analysis, in the knowledge of what actually occurred rather than what is forecast to happen. (Treasury 2004: 45)

The Green Book also states that evaluation should include ‘an assessment, quantified where possible, of what happened’ (Treasury 2004: 46) and that ‘objectives, outcomes and outputs should be defined and quantified as precisely as possible’ (ibid). Supporting the arguments of Sullivan and others, that, despite the new language of qualitative, theory-based approaches under New Labour, the UK government’s interpretation of evaluation research remained resolutely focused on quantitative, econometric assessments.

Post-2010, Conservative-led governments have been much more openly sceptical of the advantageous nature of EBP and micro-level programme evaluation. Their comments on policy evaluation are a lot more critical than the preceding Labour administrations and the assumption that ‘social science research has been able to contribute little to the normative project’ of governing (Skelcher 2008: 41) more widespread. This was outlined in the National Audit Office 2013 report *Evaluation in Government* which questioned both the usefulness of the Labour preference to micro-level evaluation to governing and the quality of the evaluations which had resulted from the positioning of EBPM as the doctrine of policymaking. The NAO was critical of the fragmented nature of both evaluation practices and evidence across government departments. It recommended a central evaluation body which presented annual compendia of evaluation research conducted on and on behalf of public bodies and government departments. The central evaluation body would have responsibilities for ensuring consistency in approaches to evaluation (including commissioning of external evaluators), resolving the problem of evaluative and information silos.

The central provocation of the 2013 NAO report was that the rhetoric of EBP had not led to a cohesive culture of policymaking through evaluation research. Echoing Sullivan (2011) and Pollitt (2014), the

report argued that the reality of public management did not match the statements by government representatives about their avowal to an ethos of EBP (Blunkett 2000). Despite the scale of investment in evaluation research—between 2010 and 2011 £44 million was spent on evaluation research by government—the *use* of evaluation research by government was limited. For example, of the 261 impact assessments of programmes and policies published in 2009–2010, only 40 made reference to evaluation evidence (NAO 2013: 8) and ‘there was little systematic evidence from government on how it has used the evaluation evidence it commissioned and produced’ (NAO 2013: 8).

As part of the 2013 report, the NAO questioned three projects of the EBPM movement: that all policy and programmes are legitimised and accompanied by evaluation research; that evaluation of policy and programmes is systematic; and the evaluation of policy and programmes is sufficiently probative to ensure ‘best practice’ in policymaking and programme working. In opposition to these claims, the report found:

Coverage of evaluation evidence is incomplete, and the rationale for what the government evaluates is unclear. Evaluations are often not robust enough to reliably identify the impact [of policies and programmes], and the government fails to use effectively the learning from these evaluations to improve impact and cost-effectiveness. (NAO 2013: 10)

That said, the report found some examples of good practice at a departmental level. In particular it praised the Department for International Development (DfID) and Department for Education, each of whom, according to the report, demonstrated the kind of systematic research for advancement in governing and administration that evaluation science claims to provide.

Post-2010 Conservative government’s position on evaluation is significantly different to New Labour’s pragmatic politics of EBPM. Evaluators who criticise or recommend changes to government policy have been labelled ‘sock puppets’ of lobby groups (Gov.uk 2015) and, where their research is funded by government as part of an EBPM-cycle, symptomatic of the ‘farce of government lobbying government’ (Hancock and Cabinet Office 2016). Under the Conservatives, policy evaluation in the United Kingdom has seen a more overt emphasis on VFM studies and cost-effectiveness (NAO 2013). Policy evaluation has been oriented towards impact and economic evaluation above process evaluation and there has

been a promotion of performance management studies and techniques. Furthermore, there has been a significant drop in the amount of resources available for evaluation research; the evaluation budget of government departments and public bodies was cut by £3 million between 2010 and 2013 (NAO 2013). Further, with the closure of the Audit Commission, the figurehead for independent evaluation of government was removed (Gris Legorreta 2015, unpublished thesis). At a doctrinal level, Conservative governments have rolled back what advances in embedding systematic, critical evaluation in the operations of government took place under New Labour through openly dismissing the practical applications of critical analysis or the appropriateness of engaging with evaluation research. Under the auspices of cost-saving and cutting bureaucracy, government departments such as the Cabinet Office have discredited the EBPM project of providing ‘good evidence’ for improving policy through critical scientific analysis ‘unsullied by ideology or value considerations’ (Botterill and Hindmoor 2012: 367) as an academic luxury that government can little afford (Hancock and Cabinet Office 2016).

Looking at youth work and the treatment of evaluation under successive governments, we are left with two images of the function and objectives of policy evaluation. On the one hand, evaluation is positioned as a technology of EBPM, directed at producing information for political decision-making. Such an understanding is resonant with how evaluation is described by evaluation science, particularly in the writing of ‘classical theorists’ of evaluation, and in commentaries on the operations of contemporary governing elites. Recently this narrative also underpins the ‘truth vs. post-truth’ debate. On the other hand, a review of how evaluation has been approached in practice and attitudes towards evaluation demonstrated by policymakers returns a decidedly different image of policy evaluation. Within this model, evaluation is a mechanism for assisting the management and organisation of policy subjects.

## A GENEALOGY OF UK EVALUATION

In genealogical terms, framing policy evaluation as a combination method focused on recording and managing rather than the production of policy through scientific research suggests that the discourse of evidence and governing is underpinned by very different discourses than the EBPM vs. post-truth conversation implies. The ‘post-truth’ accusation that the new politics of governing is characterised by the rejection

of—or simply remaining intentionally ignorant to—‘facts’ constructs contemporary politics as underscored by an adherence to rationalism and political arithmetic. The positioning of evaluation in policy debates and governing is symptomatic of the steady, continual movement towards technico-rationalist political action. However, the suggestion that evaluation is in reality more concerned with communicating information and managing policy subjects—reflected in evaluations of the youth work sector and the Youth Service—brings the role of evidence in the policy cycle much closer to discourses of NPM and information communication. Indeed, I would argue that it is these discourses which have propelled evaluation to its position of prominence in policymaking discussions. To illustrate this, I want to briefly recount each of these discourses and indicate how they have influenced the discourse of policy evaluation.

### *New Public Management*

NPM has dominated policy and policymaking since the late 1970s and became central to the approaches of the Conservative governments under Thatcher and Major. NPM also informed the ‘modernisation’ agenda of New Labour in the form of the New Public Governance (Stoker and Wilson 2004; Newman 2001, 2002, 2005; Clarke 2004a, b). While post-2010 Conservative-led administrations have altered the original frameworks of the Thatcher period—what Dunleavy et al. (2006a, b) describes as Digital Era Governance—the core components of NPM (namely agentification, outcomes-orientation, and networked governance) continue to be a prominent feature of public administration and public policy. At this point so much has been written on the emergence of NPM, its political context (specifically 1980s Thatcherism and, in the first half of the 1990s, the Major administration), and its impact on the discourse of governing (e.g. the move from government to governance, the emergence of New Public Governance), it is hardly worth devoting a great deal of space tracing its trajectory. There is already an extensive body of literature analysing the nuances of this discourse. In the light of this, I will keep my explanation of this discourse as brief as possible, focusing instead on the impact of NPM/NPG on the discourse of policy evaluation in the United Kingdom.

Central to NPM was, to use Snape’s summary, ‘the shift from government to governance, from direct service provision to enabling and competition, the re-focusing on the “customer” and “user”’ (Snape 2004:

63). According to its early proponents in the Thatcher governments of the 1980s, NPM was a resolution to the problems of centralised, ‘big’ government—namely inefficiency, gigantism and slow economic growth—through the application of the private sector ethos of economy, efficiency and effectiveness, and increasing the role of managers. NPM represented a fundamental shift in understandings of the role of government and of representatives of governing. Government was intended to facilitate the operation of the public sphere according to the principles of free market economic rationality. Light-touch regulation, empowered management regimes and the opening up of previously nationalised, government run services to a competitive market involving the private sector were all promoted as the best mechanisms for enabling service-provision which were efficient, effective and of economic VFM (the ‘3 Es’).

A marked doctrinal and operational difference in NPM and its immediate predecessor—progressive public administration or PPA (Hood 1995)—was the agentic power it assigned to top-level administrators. This is most obviously reflected in its use of the term ‘public managers’ as opposed to ‘civil servants’ or ‘public administrators’. Whereas PPA had emphasised diffuse regimes of power and control, operating under the assumption that politicians were principally driven by self-interest and could not be trusted to act in the public benefit, NPM emphasised discretionary power and freedom at managerial level. NPM criticised the ‘Jesuitical corps’ (Webb in Barker 1984: 34) of PPA who controlled the flow of and developments in social and public programme working through ‘an elaborate structure of procedural rules’ (Hood 1995: 94). Entrepreneurialism and innovation were best achieved, NPM argued, by the introduction of ‘hands on’ management styles with ‘visibly top level managers wielding discretionary power’ (Hood 1995: 97).

NPM redesigned public administration and governing according to private sector understandings of organisation management which emphasises the role of intermediaries in coordinating programmes—a *corporatised* approach or ‘networked governance’ (Newman 2001)—and ‘optimising’ outputs. Projecting itself as a ‘practical politics’ (Hunter 2015) of encouraging the public sector to be more ‘business-like’ (Clarke 2004a, b) in its operations, NPM argued that good governance was best achieved through empowering managers, while at the same time making their power contingent on the efficacy and ‘value for money’ (VFM). This practical politics, as Hunter (2015) writes:

...is pitted in supposed contrast to the realm of social and cultural differences which is positioned as qualitatively different in kind, related to issues of (private) ideological choice and preference *rather than* general human necessity. (Hunter 2015: 9, emphasis in original)

NPM constructed itself as a ‘nonpolitics’ (Duggan 2002: 10) or a ‘neutral technical policy apparatus’ (Hunter 2015: 13) where achieving positive outcomes was the priority of governing elites. While originally positive outcomes were defined solely in monetary terms (Roberts and Pollitt 1994), this style of thinking was expanded under New Labour to include non-monetised targets such as greater social and individual well-being and the enigmatic ‘community cohesion’ (Jones 2014; Stoker 2010; Newman 2005). Additionally, unlike Thatcherist NPM, New Labour emphasised a more disparate network of actors rather than corporatised ‘businesslike’ structures. Rather than empowered, ‘hands on’ managers, services would be designed and delivered by collections of stakeholders from a range of occupations and backgrounds (including members of the public, private sector organisations, public sector workers and ‘third sector’ organisations). This was expressed in their 1999 white paper *Modernising Government*. Governing power would be negotiated between and diffused among these stakeholders—an approach that became described as governance or, emphasising the links with NPM, New Public Governance (NPG). Governance, as Newman (2005), suggested that:

...the power of the nation state has been eroded and that governance processes are now characterised by multi-level interactions between different tiers within and beyond the nation state. These network based forms of coordination, it is argued, are displacing hierarchy and markets as the dominant mode of interaction. (Newman 2005: 3; see also: Kooiman 1993; Pierre and Peters 2000)

Since the return of Conservative-led administrations following the 2010 election, further shifts in governance have taken place. The vast network of organisations instituted under New Labour has been dismantled under a programme of *disintermediarisation* (Dunleavy et al. 2006a, b) and politics of austerity. This approach falls somewhere between Thatcherist NPM and New Labour NPG; there is still an emphasis on divestment of government-run services to alternative providers (including the private

sector), a prioritisation of the ‘3 e’s’, a focus on entrepreneurialism and creativity, and an orientation towards outcomes. However, under Conservative restructuring of education in particular, although *delivery* is decentralised, *management* has been steadily returned to the centre. This ‘recentralisation’ has been achieved through a reduction in budgets at a local and regional level, limiting authorities and organisations ‘freedom to manage’ (Newman 2005; Clarke 2004a, b) and restricting what they can feasibly do and, in primary and secondary education, introduction of more circumscribed curricula. Dunleavy (in Stoker et al. 2013) addresses this form of governing as Digital Era Governing (DEG).

That said, like previous administrations (Conservative- and Labour-led), governing elites since 2010 have espoused the NPM ideologies of ‘business-like’, user-centred provision and best VFM. This is articulated in the narrative of ‘choice and voice’ and initiatives such as the 2013 Choice Charter which claimed that the then coalition government would ensure that the ‘voice’ and experience of service users was paramount in the government’s strategy for public services.

NPM, NPG and DEG have had a profound impact on how policy evaluation is approached in the United Kingdom. In stark contrast to attempts by evaluation theory and science to become more nuanced, process-oriented, responsive and counterhegemonic (see for example: Guba and Lincoln 1989; Cousins and Whitmore 1998; Pawson 2013), the coexistent governing prioritisation of ‘steering, not rowing’ (Bryson et al. 2014; Newman 2001, 2005; Jones 2014) has led to evaluation become absorbed as part of a system of networked regulation, part of the audit society (Power 1997). Clarke (2004a, b) accounts for this as the *performance-evaluation nexus*, whereby policy evaluation and inspection mechanisms (e.g. quality assurance assessments, audits, reviews) are used to control the direction of actors (individual and collective) within a fragmented governance arrangement:

[T]he performance-evaluation nexus represents a solution to the problems of managing a dispersed and fragmented system ‘at arm’s length’. Withdrawing from the direct provision of services, and from the direct control of the organizations involved in provision, created new problems about how control might be exercised through other means. A rich diversity of mechanisms has emerged—contracts, commissioning, internal trading, market dynamics, partnerships, targets, outcome measurement and, of course, ‘more and better management’. (Clarke 2004a, b: 133)

According to Bjørnholt and Larsen (2014) and Gris Legorreta (2015, unpublished thesis), the influence of NPM on the shape of evaluation is most obviously reflected in the emphasis on performance indicators. The reduction in policy and programme evaluation to unambiguous measurements (Jackson 1995) is essential to, and legitimised by, the claim to ideological neutrality and ‘non-politics’ from which NPM approaches derive much of their power (see Hunter 2015 on the NPM ‘There Is No Alternative’ narrative). As Bjørnholt and Larsen write:

In the political and managerial system, performance measurement is legitimized by rhetoric referring to a more effective administration and service, increased efficiency, and more evidence-based policies. Politics become described as a technical and apolitical activity. (Bjørnholt and Larsen 2014: 405)

The ascendance of NPM in the UK policy sphere has led to evaluation becoming overtly oriented towards performance measurement (Talbot 2010). A key representation of this is the prioritisation of targets in Public Service Agreements, the main monitoring and evaluation tool for government departments, ministries and publicly funded organisations. The promotion of targets has been described by commentators on policy management and public service administration in the United Kingdom as ‘target and terror’ (Christopher and Hood 2006). The focus on performance measurement is also indicated by the definition of evaluation’s role provided by guidance manuals such as the Green Book (2004, 2011) and the Magenta Book (2006, 2011). Both of these constructs provide performance results and ensure positive performance as the central projects of evaluation research. As the Magenta Book states:

Good evaluation, and the reliable evidence it can generate, **provides direct benefits in terms of policy performance and effectiveness**, but is also fundamental to the principles of good government, supports democratic accountability and **is key to achieving appropriate returns for taxpayers’ resources**. (Treasury 2004, 2011: 12; emphasis added)

Finally, the concern with performance measurement manifests in attitudes towards evidence in evaluation research. Despite the wealth of literature on the usefulness and accuracy of qualitative, process-oriented evidence (Guba and Lincoln 1981; Lincoln and Guba 1986; Denzin and

Giardina 2008), guidance on and reviews of evaluation research in the United Kingdom have prioritised ‘unambiguous’ performance measures (Jackson 1995) in particular quantitative evidence produced through economic and positivist approaches. This has been labelled by St. Pierre (2002) and House and Howe (2001) as reflective of ‘methodological fundamentalism’ and by Torrance (2006) as illustrative of a systemic promotion of ‘scientism’. In the United Kingdom, guidance documents such as the Green Book, large-scale evaluations such as the evaluation of the Cancer Strategy and critical reviews on the use of evaluation all demonstrate a preference for measurement-oriented evidence (comparative impact evidence, cost effectiveness evidence, outcome evidence).

### *ICT/Informational Theory*

If the orientation of evaluation towards management is representative of the growth of NPM, then its prioritisation of the distillation of evidence and articulation of the shape of evaluands—including the Youth Services—is a reflection of the influence of growing debates around and developments in information communication technology (ICT). ICT is principally concerned with the transmission—and dynamics of the transmission—of information packages in order to make ever-expanding information terrains manageable. Whereas the locus of cultural theoretical positions is the *content* of information, in terms of discourse, representations or semiotics, the interest of informational theory is the form and flow of information conceptualised as a material object. Traditionally this came under the heading of documentation and cybernetics but more recently this has been framed in terms of Big Data or datafication debates. In the context of policy evaluation, while the influence of NPM has been well documented, the importance and impact of ICT on the emergence of evaluation as a feature of policymaking has to date been under recognised. In the light of this, I want to spend some time outlining the ‘information debate’.

In his 2008 work *Modern Invention of Information: Discourse, History and Power*, Day traces the contemporary debate around the communication and interpretation of information within information-rich environments back to emergence of documentation as a profession at the end of the nineteenth century. Spearheaded by the writing of Paul Otlet, Henri Lafontaine (who later received the Nobel Peace Prize) and Suzanne Briet’s documentation aimed to reconstruct books and other

‘informational objects’ (Day 2008: 10) as machines which enable the movement and storage of packages of ‘knowledge’ in the form of facts. This is articulated in Otlet’s seminal contribution *Traité de documentation* (1934) in which he uses the trope of ‘the book-machine’ to advance a conceptualisation of informational objects (principally paper documents) as both containing and transmitting thought.

As a field of practice, documentation distinguished itself from affiliated disciplines—specifically librarianship—but focusing on the technical retrieval, cataloguing and distribution of material representations of knowledge. Proponents such as Otlet and Briet applied systems approaches to interpreting the nature and function of documents within modernity. This involved a consideration of how documentation facilitated—or not—the machinations of the social system and a development of ideas for how documentation could assist the social system better. The rationale behind documentation advocates concern was the pressing question of how to manage—and utilise—the sheer amount of documentary materials available to scientists effectively. As Briet (1951) argued, the availability of extensive resources in libraries had not translated into more scientific advancements:

The cumulative documentation at the disposal of the human sciences overwhelms in importance and in quantity the figures, however impressive, of scientific production, per se. [...] The immense libraries with which the scholar surrounds himself, and those which he consults beyond his abode, are for him an exploration, partly untapped. (Briet 1951: 14)

Briet presented her concerns as the inevitable outcome of the industrialisation and globalisation which, for her, typified modernity. She argued that as part of the expansion of capitalist modes of production, intellectual labour had become increasingly industrialised and oriented towards the production of information objects or documents. As a result, a global problem of overproduction had occurred (Day 2008).

The resolution to this, for Briet and others, was the development of techniques and technologies for processing the contents of documentary materials systematically. In doing so, potential users would be better able to access documents relevant to their field of study or analysis. Of critical importance to documentary scientists were what Otlet (1934) labelled *réseau* (webs) between informational objects or ‘the relationship of books to other books’ (Day 2008: 14). Of paramount importance

was the development of a means of connecting the vast amount of documentary materials available globally so that they were accessible for consultation and analysis by individual users. As Day (2008) describes, the principal aim of documentation, as laid out by Otlet (1934):

[is] the creation of a technological device that would unify information but also transform it in such a way as to present it in the most ‘advantageous’ manner to each viewer. (Day 2008: 19)

As outlined in works such as Otlet’s *Traité* and Briet’s *Qu’est-ce que la documentation?* (1951), documentation positioned documentary materials or informational objects as *technologies* which facilitated the development and communication of discrete amounts of knowledge through acting as both ‘materially fixed’ and being ‘susceptible of being used for consultation, study, and proof’ (Briet 1951: 7). Applied in the contemporary jargon of information, documentation presented materials as interfaces between users and facts. Moreover, unlike librarianship, documentation emphasised the productive nature of this interaction and the function of documentary materials to extend knowledge and aid social advancement. Informational objects do not just assist in the cataloguing of canonical knowledge but enable the production of new knowledges. In this sense they are *dynamic* and can be used to gather and develop knowledge. Otlet (1934) discusses this using the term *répétition amplifiante* (literally amplifying repetition). Books repeat what came before and are used to extend what came before. As Day summarises:

Repetition, as amplification, leads to the universal and ‘geometric’ expansion of knowledge. [...] For Otlet, texts are both vehicles and embodiments of dynamic repetition, leading to an expansion of knowledge and a change in the form of knowledge. (Day 2008: 14)

Documentation’s position as the originator of information science is perhaps most obviously represented by the rhetoric of technology employed by authors such as Otlet and Briet. However, within Day’s history, documentation makes five central contributions which affirm its status as the beginning of the discourse of information science. These are: (i) a concern with interfaces and shared meanings; (ii) a recognition of the dynamic and productive relationship between informational objects and users; (iii) an awareness of the overproduction of informational objects;

(iv) an imagining of the use of informational objects as involving a feedback loop; and (v) an awareness of the importance of the ordering of objects and networks between informational objects.

While documentation is positioned by Day as the earliest discussion into the need to transmit and connect vast arrays of information and how best this could be done, the foundational text in informational studies/ICT as a discourse or discipline according to historical analyses of information theory (Day 2008; Kline 2006; Terranova 2004) is Claude Shannon and Warren Weaver's 1949 *The Mathematical Theory of Communication*. Based on a solo-authored work of Shannon's with the same name, this work provides a model of information directed by communication engineering. In this work, the authors present what Reddy and Day have addressed as the 'conduit metaphor'. As Day summarises:

According to this model or metaphor, information is the flow and exchange of a message, originating from one speak, mind, or source and received by another. (Day 2008: 38)

According to Terranova (2004), Shannon progresses from the following fundamental position:

- Proposition I: *Information* is what stands out from *noise*  
 Corollary I(a): Within informational cultures, the struggle over *meanings* is subordinated to that over '*media effects*'  
 Corollary I(b): The cultural politics of information involves a return to the *minimum conditions of communication* (the relation of *signal* to *noise* and the problem of *making contact*)

Terranova 2004: 10 (emphases added)

Analysing information using a communication engineering/ICT framework, Shannon states that, for communicative devices, 'two messages, one of which is heavily loaded with meaning and the other of which is pure nonsense, can be exactly equivalent' (Shannon, citation from Terranova 2004: 12–13). Shannon argues that the principle concerns of information theory are threefold: (i) the clearing of a channel between sender and receiver; (ii) the transmission of a message or signal from sender and receiver; and (iii) the ability of the recipient to distinguish a signal from other 'media effects' (noise). When Shannon was writing,

however, these concerns were not associated with information theory or ICT but with the emergent interdisciplinary field of cybernetics (Ashby 1956; Wiener 1948).

According to Shannon and members of the cybernetics movement, information was defined by its status as the object of communication between actors. What distinguished information from noise was that it was sent along a channel for a recipient to interpret. Importantly here, it is the intent of the sender and the ability of the recipient to receive and interpret the message which defines something as information. The focus of technologists should therefore be the stability of the channel or pathway between sender and receiver. This, as Terranova writes:

[A]ttributes a *secondary* importance to the question of the *meaning* of messages when compared to the basic problem of *how to increase the effectiveness of the channel*. (Terranova 2004: 14)

The concern of elites, from this perspective, is twofold: reinforcing the connection between sender and receiver and damping the sounds outside the information channel. The content and meaning of the message matter less than excluding interference or ‘holding off noise’ (ibid: 15). Persuasion and convincingness are secondary to the strength of connection with receivers and silencing of alternate interpretations and messages. To suggest the dominance of these concerns in contemporary politics, Terranova uses the example of televised political debates suggesting that:

In this context, the opponent becomes noise and the public becomes a target of communication: not a rational ensemble of free thinking individuals, endowed with reason, who must be persuaded, but a collective receiver to which a message can be sent only on condition that the channel is kept free of noise (competing politicians, but also the whole noisy communication environment to which such politicians relate, where, for example, more young people vote for reality TV shows than for general elections). (Terranova 2004: 17)

At present, this process of ‘clearing out’ noise does not solely involve technical advancement but governance of attention (Crogan and Kinsley 2012) and damping out of alternate information. The need for governments to engage in this is made more pronounced by the expansion of

communicative technologies and opening of new channels between and across users. This is what distinguishes the ‘modern era’ from the period before the expansion of communicative technologies which operated through the transmission of signals (i.e. telephony, television and the internet). As Jordan (2015) notes:

The politics of information has always been present [...] but it has changed. The times now are characterised by an information flood driven on by the cultures of the digital and the internet. These new times have inverted hundreds of years of information scarcity. (Jordan 2015: 23)

Within this informational milieu, the process of establishing and clearing out channels becomes more complex and involves the mobilisation of ‘sophisticated forms of “viral” propaganda’ and ‘the takeover of telecom giants, venture capital and banks, and the sharp rise in regulatory efforts by governments’ (Lovink 2002: 308). These debates have coalesced under the notion of *datafication*.

Datafication, according to Lycett (2013), refers to both the ‘re-representation’ (p. 383) of the world through the algorithmic analysis of codes (as the organised manifestation of the dematerialised form of an object) and the production of a political and economic sphere directed by the movement and mobilisation of units of data without reference to their content. In many ways this is merely an updated version of cybernetics, particularly the emphasis on flow. However, what marks datafication apart from cybernetics is (i) its treatment of information overload as potentially positive if manipulated and (ii) its positioning of the process of differentiating noise from signals as affective modulation through the operation of codes (Cheney-Lippold 2011).

The origins in datafication lie in the introduction and spread of wireless and mobile technologies (particularly smartphones) and the associated ascendance of Big Data debates which focused on the rapid, exponential growth of datasets. Big Data is defined by Boyd and Crawford (2012) as ‘a cultural, technological, and scholarly phenomenon that rests on the interplay of:

1. *Technology*: maximizing computation power and algorithmic accuracy to gather, analyze, link, and compare large data sets
2. *Analysis*: drawing on large data sets to identify patterns in order to make economic, social, technical, and legal claims.

3. *Mythology*: the widespread belief that large data sets offer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with the aura of truth, objectivity, and accuracy'

boyd and Crawford 2012: 663

Central to the era of Big Data is the management, use and function of large—and expanding—datasets. Against this landscape of information, according to Castells (2011), the content of signals, ceases to be a concern. Instead of minimising distortion through statistical differentiation, discussions focus on increasing the intensity, speed and volume of traffic (c.f. Castells 2011; Loader 1997; Mortier et al. 2014). In the face of ever increasing content—which, thanks to technological advances, was not restricted to any particular format—the ‘data smog’ (Lovink 2002; Mortier et al. 2014) becomes unavoidable and the concern of the information complex (Castells 2011) shifts from maintaining a clear channel to increasing the speed and volume of information flow between users.

Moreover, again due to the voluminous nature of data, datafication moved away from circuit or network models of the dynamics of information transmission and user interaction, towards a discourse oriented towards the aesthetic and emotional intensities of Big Data. This is noted by both Lycett (2013) in his discussion of the role of graphics and aesthetic manipulations of large datasets and Clough (2004) whose work on ‘technoscience’ argues that the contemporary discourse of information is concerned with how flows direct ‘attention, arousal, interest, receptivity, sensation, attentiveness, action, reaction and inaction’ (Clough 2004: 15). According to these perspectives, data are transmitted at the level of subconscious intensity. Even where graphics are used, Lycett argues, they are not deployed as non-textual representations but visual evocations separated from their encoded meanings. Sellar (2015) extends this by arguing that datafication is characterised by the aim for data to be ‘felt’ before it is either separated into schematics of probable meanings (the cybernetic position) or connected with associated knowledges.

Perhaps the most notable feature of datafication is the attempt to unsettle the *need* to interpret data. This trend is particularly prevalent in debates about Big Data. As part of a general investigation into the dynamics of large datasets, how they can be used, and how they can and should be analysed, boyd and Crawford (2012) track the emergence of

the idea that, due to the scale of data, information no longer requires interpretation, either at the level of hermeneutic meaning-making or cybernetic sense-making. This claim was articulated most overtly by Anderson, Editor-in-Chief of *Wired*, when he stated:

This is a world where massive amounts of data and applied mathematics replace every other tool that might be brought to bear. Out with every theory of human behaviour, from linguistics to sociology. Forget taxonomy, ontology, and psychology. Who knows why people do what they do? The point is they do it, and we can track and measure it with unprecedented fidelity. With enough data, the numbers speak for themselves. (Anderson 2008: np; quotation from: Boyd and Crawford 2012: 666)

Notably, although boyd and Crawford (2012) present Anderson's comment as reflecting a preference towards large amounts of numerical evidence, it can also be presented as an emphasis on the look and shape of information over the nature of the evidence (i.e. quantitative or qualitative). Anderson's statement promotes a vision of 'good data' as being able to speak without interpretation. As such, the role of applied mathematics is not to translate but to allow data to have form without becoming fixed. It must be able to speak for itself. Within the landscape of Big Data and datafication, this is achieved through the use of schema. Traffic is made visible through the application of code which, through categorisation, creates 'digital walls' (Cheney-Lippold 2011: 166) around disparate, multimodal forms of data. This suggests a movement from the immaterial (flow) to the material (code) and back again (traffic). The signal is coded into 'data packages' (Castells 2011) so that it can impact on the level of immateriality (feeling, sensation) through volume. Datafication arguably represents what Castells argued would be the defining features of the information society—the separation of content from information processing and the emphasis on collating a multiplicity of digitised formats into a single flow of traffic through the application of code. Anderson's statement presents the potential of data to transmit meanings as dependent not on the possibility of interpretation but on the possibility to be loud and to be presented as a coherent block of sound.

The growing importance of information theory and ICT has contributed to the orientation of UK evaluation making the transmission of information clearly and concisely a priority. Within ICT, reducing vast

quantities of multimodal data to set transmission patterns is essential in order to avoid the problems of data smog (Lovink 2002), white noise (Shannon 1948) or entropy (Wiener 1946). Ambiguity and complex messages are to be avoided. Transposed to the context of communicating the merits of policies or programmes, ICT encourages evaluators and governing elites to focus their analysis on reductionism and simplification. The echoes of ICT science demands for clarity of transmission and interlinked data (the predominant concerns of cybernetics and documentation respectively) are visible in the promotion of unified, ‘clean’ evidence in the context of EBPM. The objective of evaluation according to this perspective is to, if not synthesise a vast array of information into a simplified singularity (Cairney 2016), at least facilitate the communication of ‘highlights’ which can be passed through the ever-expanding field of knowledge management systems with minimal distortion (Pawson 2006). The influence of ICT and information theory is also more directly reflected in the concern of the Green and Magenta Books and the 2013 NAO review of evaluation with the communication and dissemination of findings and the need to be mindful of communication in the design of evaluative approaches.

In the context of youth work, the traces of ICT are visible in the positioning of evaluative frameworks—such as the HCEC *framework for outcomes*—as mechanisms for articulating impact in discrete information packages. The framing of ‘unclear’ information as fundamentally problematic—a key feature of Shannon’s discussion of noise and Weiner’s arguments on entropy—is further reflected in the HCEC review’s attitude towards the vast amount of accounts of the positive contribution of youth work submitted by young people, stakeholders and youth workers. Rather than presenting this evidence as a clear demonstration of the merits of youth work, the committee criticised the Youth Service and youth work organisations for failing to reduce these accounts to concise data packages. This position is fundamentally at odds with contemporary evaluation science—particularly that espoused by advocates of feminist or critical approaches (Brisolara et al. 2014; Guba and Lincoln 1989)—which allows for mixed method approaches and multiple forms of evidence. However, it falls precisely within the models of ICT promoted by Shannon, Weiner and others which emphasises the need for ‘clarity’.

## A DIFFERENT DISCOURSE OF EVALUATION AND GOVERNING

Writing on the Evaluative State, Neave (1988) argues that the purpose of evaluative frameworks is not to produce evidence about good or bad practice in Higher Education (the location of his analysis) but to ensure that subjects of Higher Education Institutions (HEIs) adhere to desires of central management. The focus on ‘remote steering’ (Neave 1988: 267) and redesign of HE into a complex system of administrative and implementing parts has led to evaluation occupying a key role in the running of HE. As HEIs have become more akin to corporate systems of departments, institutes, schools and numerous layers of ‘intergovernment’ or ‘middle management’, evaluation has emerged as a technique for ensuring that the overall objectives of the university or college—as the body corporate—continue to be worked towards at optimal levels of efficiency and tangible VFM. The sole objective of Evaluative State, Neave argues, is to correct and verify behaviour and performance of individuals so that operational goals can be achieved in the best way possible. Evaluation research—embodied in either examinations, student-voice surveys or appraisals of research impact—is a mechanism for making increasingly corporatised educational environments manageable.

The review of evaluation above would support Neave’s argument that evaluation is more about management and organisation than producing evidence for good policy and practice. This is not entirely antithetical to how evaluation science presents evaluation. ToC models, for example, construct evaluation as a method of assisting programme planning and management. That said the operationalisation of evaluation in youth work and historic and contemporary attitudes to evaluation among governing elites differs from the narrative of policy evaluation in its construction of the relationship between evaluation and governing. Under evaluation science and discourses of EBPM and ToC, evaluation enables political decision-making through verifiable and scientifically legitimate evidence. However, the review presented above suggests that decision-making is disconnected from evaluative evidence. Evaluation in the United Kingdom focuses on maximising efficiency, providing a clear message about the nature of an evaluand, and ‘mapping’. The fact that evaluation has not been adopted as a means for policymaking elites to identify the merits of a particular policy approach is most clearly represented in the fact that evaluations are prohibited from making comments on the merits of policy.

What I have argued in this chapter is that the prominence and mobilisation of evaluation in the United Kingdom is not a manifestation of evidence-driven political decision-making but a reflection of emergent discourses of NPM and ICT. Evaluation is therefore oriented towards governing networked systems at a difference and simplifying the ‘message’ of evaluands. This is visible in youth work evaluations which have historically and continue to operationalise evaluation as a means to bound off what is a very disparate sector into a specific set of organising principles and core objectives. Having established managing and distilling policy subjects as the primary foci of policy evaluation research, I now want to demonstrate how this constitutes a form of disciplinary and controlling approach to governing. To illustrate this, I want to interrogate policy evaluation using critical sociological theories of governmentality (specifically the work of Foucault and Deleuze). This will be the focus of the next chapter.

## REFERENCES

- Ashby, William Ross. 1956. *An introduction to cybernetics*. New York: Wiley.
- Barker, Rodney. 1984. The Fabian state. In *Fabian essays in socialist theory*, 27–38. London: Heinemann.
- Benjamin, Lehn M., and Jennifer C. Greene. 2009. From program to network: The evaluator’s role in today’s public problem-solving environment. *American Journal of Evaluation* 30 (3): 296–309.
- Bevan, Gwyn, and Christopher Hood. 2006. Have targets improved performance in the English NHS? *BMJ* 332 (7538): 419–422.
- Bjørnholt, Bente, and Flemming Larsen. 2014. The politics of performance measurement: ‘Evaluation use as mediator for politics’. *Evaluation* 20 (4): 400–411.
- boyd, Danah, and Kate Crawford. 2012. Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society* 15 (5): 662–679.
- Botterill, Linda Courtenay, and Andrew Hindmoor. 2012. Turtles all the way down: Bounded rationality in an evidence-based age. *Policy Studies* 33 (5): 367–379.
- Briet, Suzanne. 1951. *Qu’est-ce que la documentation?* vol. 1. Paris: Éditions documentaires, industrielles et techniques.
- Brisolara, Sharon, Denise Seigart, and Saumitra SenGupta (eds.). 2014. *Feminist evaluation and research: Theory and practice*. Guilford Publications.

- Bryson, John M., Barbara C. Crosby, and Laura Bloomberg. 2014. Public value governance: Moving beyond traditional public administration and the new public management. *Public Administration Review* 74 (4): 445–456.
- Cairney, Paul. 2016. *The politics of evidence-based policy making*. London: Springer.
- Castells, Manuel. 2011. *The rise of the network society: The information age: Economy, society, and culture*. vol. 1, Oxford: John Wiley & Sons.
- Cheney-Lippold, John. 2011. A new algorithmic identity soft biopolitics and the modulation of control. *Theory, Culture & Society* 28 (6): 164–181.
- Chen, Huey-Tsyh. 1994. Theory-driven evaluations: Need, difficulties, and options. *Evaluation Practice* 15 (1): 79–82.
- Clarke, John. 2004a. Dissolving the public realm? The logics and limits of neo-liberalism. *Journal of Social Policy* 33 (01): 27–48.
- Clarke, John. 2004b. *Changing welfare, changing states: New directions in social policy*. London: Sage.
- Clough, Patricia Ticineto. 2004. Future matters: Technoscience, global politics, and cultural criticism. *Social Text* 22 (3): 1–23.
- Cousins, J. Bradley, and Elizabeth Whitmore. 1998. Framing participatory evaluation. *New Directions for Evaluation* (80): 5–23.
- Crogan, Patrick, and Samuel Kinsley. 2012. Paying attention: Toward a critique of the attention economy. *Culture Machine* 13: 1–29.
- Day, Ronald E. 2008. *The modern invention of information: Discourse, history, and power*. Carbondale, IL: SIU Press.
- Denzin, Norman K. 2009. The elephant in the living room: Or extending the conversation about the politics of evidence. *Qualitative Research* 9 (2): 139–160.
- Denzin, Norman K., and Michael D. Giardina. 2008. *Qualitative inquiry and the politics of evidence*. Walnut Creek: Left Coast Press.
- Department of Education and Science. 1982. *Experience and participation: Review Group on the Youth Service in England* (The Thompson Report). London: HMSO.
- Department of Education and Science. 1987. *Effective youth work: A report by HM inspectors, education observed*, vol. 6, London: Department of Education and Science.
- Dunleavy, Patrick, Helen Margetts, Simon Bastow, and Jane Tinkler. 2006a. New public management is dead—long live digital-era governance. *Journal of Public Administration Research and Theory* 16 (3): 467–494.
- Dunleavy, Patrick, Helen Margetts, Simon Bastow, and Jane Tinkler. 2006b. *Digital era governance: IT corporations, the state, and e-government*. Oxford: Oxford University Press.
- Fischer, Frank. 2003. *Reframing public policy: Discursive politics and deliberative practices*. Oxford: Oxford University Press.

- Gov.uk. 2015. *Eric Pickles cracks down on wasteful spending of government lobbying government*—Press releases. <https://www.gov.uk/government/news/ericpickles-cracks-down-wasteful-spending-of-government-lobbying-government>. Accessed 30 May 2016.
- Gris Legorreta, P.C. 2015. *The hidden politics of evaluation: Towards a smarter state?* PhD dissertation, University of Sheffield.
- Guba, Egon G., and Yvonna S. Lincoln. 1981. *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches*. San Francisco: Jossey-Bass.
- Guba, Egon G., and Yvonna S. Lincoln. 1989. *Fourth generation evaluation*. London: Sage.
- Hammersley, Martyn. 2013. *The myth of research-based policy and practice*. London: Sage.
- Hancock, Rt. Hon. Matthew, and Cabinet Office. 2016. *Government announces new clause to be inserted into grant agreements*. London: Cabinet Office Press Office. <https://www.gov.uk/government/news/government-announces-new-clause-to-be-inserted-into-grant-agreements>. Accessed 01 June 2017.
- Hood, Christopher. 1995. Emerging issues in public administration. *Public Administration* 73 (1): 165–183.
- Hood, Christopher. 2006. Gaming in targetworld: The targets approach to managing British public services. *Public Administration Review* 66 (4): 515–521.
- House of Commons Education Committee. 2011. *Services for young people: Third report of the session 2010–12*, vol. 1, HC 755-1. London: HM Stationery Office.
- Hunter, Shona. 2015. *Power, politics and the emotions: Impossible governance?* Abingdon: Routledge.
- Jackson, Peter. 1995. Reflections on performance measurement in public service organizations. In *Measures for success in the public sector, Public Finance Foundation*, 1–18. London: CIPFA.
- Jenkins, Bill, and Andrew Gray. 1990. Policy evaluation in British government: From idealism to realism. In *Program evaluation and the management of government: Patterns and prospects across eight nations*, 53–70. New Brunswick, NJ: Transaction Publishers.
- Jones, Rhys, Jessica Pykett, and Mark Whitehead. 2014. The geographies of policy translation: How nudge became the default policy option. *Environment and Planning C: Government and Policy* 32 (1): 54–69.
- Jordan, Tim. 2015. *Information politics: Liberation and exploitation in the digital era*. London: Pluto Press.
- Kippin, Henry, Gerry Stoker, and Simon Griffiths (eds.). 2013. *Public services: A new reform agenda*. London: Bloomsbury.
- Kline, Ronald R. 2006. Cybernetics, management science, and technology policy: The emergence of “information technology” as a keyword, 1948–1985. *Technology and Culture* 47 (3): 513–535.

- Kooiman, Jan (ed.). 1993. *Modern governance: New government-society interactions*. London: Sage.
- Lincoln, Yvonna S., and Egon G. Guba. 1986. But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. *New Directions for Program Evaluation* 1986 (30): 73–84.
- Loader, Brian. 1997. *The governance of cyberspace: Politics, technology and global restructuring*. Hove: Psychology Press.
- Lovink, Geert. 2002. *Dark fiber: Tracking critical internet culture*. Cambridge, MA: MIT Press.
- Lycett, Mark. 2013. ‘Datafication’: Making sense of (big) data in a complex world. *European Journal of Information Systems* 22 (4): 381.
- MacDonald, Barry. 1976. Evaluation and the control of education. In *Evaluation: The state of the art*, ed. D. Tawney, 125–136. London: Schools Council.
- Mertens, Donna M. 1999. Inclusive evaluation: Implications of transformative theory for evaluation. *American Journal of Evaluation* 20 (1): 1–14.
- Milson, Fred. 1970. *Youth work in the 1970s*. London: Routledge/Thoemms Press.
- Ministry of Education. 1960. *The Youth Service in England and Wales* (The Albemarle Report). London: HMSO.
- Mortier, Richard, Hamed Haddadi, Tristan Henderson, Derek McAuley, and Jon Crowcroft. 2014. Human-data interaction: The human face of the data-driven society. SSRN. <https://ssrn.com/abstract=2508051>.
- Neave, Guy. 1988. On the cultivation of quality, efficiency and enterprise: an overview of recent trends in higher education in Western Europe, 1986–1988. *European Journal of Education* 7–23.
- Neave, Guy. 1998. The evaluative state reconsidered. *European Journal of Education* 33 (3): 265–284.
- Newman, Janet. 2001. *Modernizing governance: New Labour, policy and society*. London: Sage.
- Newman, Janet. 2002. The new public management, modernization and institutional change. In *New public management: Current trends and future prospects*, 77–91. London: Routledge.
- Newman, Janet. 2005. *Remaking governance: Peoples, politics and the public sphere*. Bristol: Policy Press.
- Ord, Jon. 2014. Aristotle’s phronesis and youth work: Beyond instrumentality. *Youth & Policy* 112: 56–73.
- Otlet, Paul. 1934. *Traité de documentation: le livre sur le livre, théorie et pratique*. Bruxelles: Editiones Mundaneum.
- Pawson, Ray. 2006. *Evidence-based policy: A realist perspective*. London: Sage.
- Pawson, Ray. 2013. *The science of evaluation: A realist manifesto*. London: Sage.
- Pawson, Ray, and Nick Tilley. 1997. *Realistic evaluation*. London: Sage.

- Peters, B. Guy, and Jon Pierre. 2000. Citizens versus the new public manager: The problem of mutual empowerment. *Administration & Society* 32 (1): 9–28.
- Pollitt, Christopher. 1993. Occasional excursions: A brief history of policy evaluation in the UK. *Parliamentary Affairs* 46 (3): 353–363.
- Pollitt, Christopher. 2014. Forty years of public management reform in UK central government: Promises, promises.... In *Rethinking policy and politics: Reflections on contemporary debates in policy studies*, ed. S. Ayres, 7–28. Bristol: Policy Press.
- Power, Michael. 1997. *The audit society: Rituals of verification*. Oxford: Oxford University Press.
- Reynolds, Shirley, and Elizabeth Trinder. 2008. *Evidence-based practice: A critical appraisal*. Oxford: John Wiley & Sons.
- Roberts, Simon, and Christopher Pollitt. 1994. Audit or evaluation? A national audit office VFM study. *Public Administration* 72 (4): 527–549.
- Ryle, Michael, and Peter G. Richards (eds.). 1988. *The commons under scrutiny*. Abingdon: Taylor & Francis.
- Sanderson, Ian. 2000. Evaluation in complex policy systems. *Evaluation* 6 (4): 433–454.
- Scriven, Michael. 1981. *The logic of evaluation*. Inverness, CA: Edgepress.
- Sellar, Sam. 2015. A feel for numbers: Affect, data and education policy. *Critical Studies in Education* 56 (1): 131–146.
- Skelcher, Chris. 2008. Does governance perform? Concepts, evidence, causalities, and research strategies. In *Managing to Improve Public Services*, ed. Jean Hartley, Cam Donaldson, Skelcher Chris, and Mike Wallace, 3–23. Cambridge: Cambridge University Press.
- Snape, Stephanie. 2004. Liberated or lost souls. Is there a role for non-executive councillors?. In *British local government into the 21st century*, ed. Gerry Stoker and David Wilson, 60–75. Basingstoke: Palgrave Macmillan.
- Squires, Peter, and Lynda Measor. 2005. Below decks on the youth justice flagship: The politics of evaluation. In *The politics of evaluation: Participation and policy implementation*, ed. David Taylor and Susan Balloch, 21–40. Bristol: Policy Press.
- St. Pierre, Elizabeth Adams. 2002. ‘Science’ rejects postmodernism. *Educational Researcher* 31 (8): 25–27.
- Stoker, Gerry, and Alice Moseley. 2010. Motivation, behaviour and the micro-foundations of public services. *London* 2020: 23.
- Stoker, Gerry, and David Wilson. 2004. *British local government into the 21st century*. Basingstoke: Palgrave Macmillan.
- Stone, Deborah A. 2002. *Policy paradox: The art of political decision making*. New York: Norton.

- Sullivan, Helen. 2011. 'Truth'junkies: using evaluation in UK public policy. *Policy & Politics* 39 (4): 499–512.
- Talbot, Colin. 2010. *Performance in government: The evolving system of performance and evaluation measurement, monitoring, and management in the United Kingdom*. Washington, DC: World Bank. Independent Evaluation Group (IEG).
- Taylor, David, and Susan Balloch. 2005. *The politics of evaluation: Participation and policy implementation*. Bristol: Policy Press.
- Terranova, Tiziana. 2004. *Network culture: Politics for the information age*. London: Pluto Press.
- Torrance, Harry. 2008. Building confidence in qualitative research: Engaging the demands of policy. *Qualitative Inquiry* 14 (4): 507–527.
- Treasury, H.M. 1988. *Policy evaluation: A guide for managers*. London: The Stationery Office.
- Treasury, H.M. 2004. *The Green Book*. London: The Stationery Office.
- Treasury, H.M. 2011. *The Magenta Book*. London: The Stationery Office.
- Weiss, Carol H. 1972. *Evaluation*. Upper Saddle River: Prentice-Hall International.
- Wiener, Norbert. 1948. *Cybernetics*. Paris: Hermann.
- Young, Ken, Deborah Ashby, Annette Boaz, and Lesley Grayson. 2002. Social science and the evidence-based policy movement. *Social, Policy and Society* 1 (3): 215–224.
- Young Foundation. 2012. *An outcomes framework for young people's services*. Available at: <http://www.youngfoundation.org/publications/reports/an-outcomes-framework-young-peoples-services>.



<http://www.springer.com/978-1-137-54512-1>

Evaluation and Governing in the 21st Century  
Disciplinary Measures, Transformative Possibilities

Duffy, D.N.

2017, XV, 161 p., Hardcover

ISBN: 978-1-137-54512-1

A product of Palgrave Macmillan UK