Jnl Soc. Pol., 34, 4, 601–618 © 2005 Cambridge University Press
doi:10.1017/S0047279405009177 Printed in the United Kingdom

Governing Through Evidence: Participation and Power in Policy Evaluation

DAVID TAYLOR

School of Applied Social Science, University of Brighton, Falmer, Brighton BN1 9PH, email: d.taylor@bton.ac.uk

Abstract

This article considers different approaches to policy evaluation within recent writings on governance and evaluation research. Governance theorists locate evaluation within a new 'mode of regulation' aimed at managing a 'dispersed state at arms length'. The strength of this approach is that it places policy evaluation within an explicitly *political framework*. However, the governance perspective focuses on evaluation as a state-centred activity and tends to underplay its *contested* nature. Within the evaluation research literature, on the other hand, attempts to deal with the 'politics of evaluation' tend to focus on evaluation *as a research practice*. This misses the discursive construction of evaluation as a political project. While some approaches within the evaluation literature stress the 'democratising' role of 'stakeholder participation', they are set within a democratic pluralist model of politics, which frequently fails to deal with underlying power relations within which the clash of competing stakeholder interests takes place. This article challenges these constructions. It draws on the strengths of both approaches to propose a critical 'politics of evaluation' based on service user control. It suggests some practical strategies drawn from a review of recent evaluation research.

Introduction

Recent years have seen an 'evidential turn' in the political language of public policy. This has generated increased attention to policy and programme evaluation. In terms of a wider analysis, this can be seen in the literature on governance and social policy (see Clarke, 2003, 2004; Clarke *et al.*, 2000; Daly, 2003; Glendinning *et al.*, 2002; Newman, 2002). A key strength of this work is the way it contextualises evaluation, locating it within changing forms of governance. At the most general level, evaluation is seen as a socially constructed and politically articulated practice. The increasing role of evaluation as a political phenomenon can be traced in Britain to the rise of neo-liberalism, public choice theory and new managerialism over the last two decades of the twentieth century. These movements created a drive towards 'evidence' as a basis for consumer choice. Early expressions of this formulation at the political level can be found in the rise of the citizens' charter idea, initially promoted by the John Major government and subsequently adopted across the British political spectrum in the early 1990s (see

https://doi.org/10.1017/S0047279405009177 Published online by Cambridge University Press

Taylor, 1992). The evolution of this approach led to Tony Blair's now famous claim that 'what counts is what works' (cited in Clarke, 2003: 1) and the consolidation of evidence-based policy and practice at the heart of the New Labour policy agenda.

Policy evaluation is not new, of course (see Hogwood and Gunn, 1984). The measurement of policy outcomes and programme outputs has long been seen as central to social policy as an academic discipline and an administrative practice. The historical focus within academic approaches has been on developing appropriate means of measurement to secure 'hard evidence'. It has been developed as a technique, whose role is to pin down the ever-changing world of service delivery. It is, on this view, *instrumental* and *contingent*: 'if we lived in a world of complete certainty and perfect administration there would be no need for evaluation' (Hogwood and Gunn, 1984: 219). Something different, however, has happened in recent times. From being a contingent instrument of administration, evaluation has become a central legitimating device for a new form of politics. A political critique emerged from within government in the 1990s, which saw much academic research as driven by dogma. Political values should be stripped out in favour of 'objective scientific' evidence of 'what really works'. Along with audit and performance management, evaluation has become a key entry in the lexicon of new public management and the pursuit of what politicians increasingly frame as 'value-free policy'.

At the same time, evaluation research (some commentators see the term 'research' as what distinguishes evaluation from audit) is a well-established independent scientific practice with its own body of theory. Evaluators and social researchers have adopted a range of epistemological standpoints. Despite some bitter 'paradigm wars' within evaluation theory between realists and social constructionists, there is now a broader range of theory available to evaluators. Social constructionism has moved beyond a simple relativism towards discursive analyses of power, and critical realists have developed post-positivist positions that are no longer universalistic (see, for example, Bhaskar, 1975; Archer, 1995; Craib, 1997; Schilling, 1997). Despite these developments, there remain fundamental ontological and epistemological disagreements, which divide evaluation researchers both in their theoretical orientation and, more importantly, in their understanding of the political realm. Yet across these divisions there is a common acknowledgement that evaluation is political. What is at issue here is the differential and frequently partial understanding of politics employed in many of these approaches.

Social constructionists, employing a democratic pluralist model of politics, have been most prominent in these debates (see, for example, Gubba and Lincoln, 1989, and, more recently, Fetterman, 2001), especially in their critiques of the realist pursuit of scientific objectivity. 'To approach evaluation scientifically is to miss completely its fundamental social, political and value-oriented character'

https://doi.org/10.1017/S0047279405009177 Published online by Cambridge University Press

(Gubba and Lincoln, 1989: 7). However, Pawson and Tilley, the prominent advocates of 'realistic evaluation' and arch enemies of constructivism's view of values as situational, also argue, 'the very act of engaging in evaluation constitutes a political statement' (1997: 11). Both employ, as I argue below, an incomplete conceptualisation of the political.

In very different ways, both the governance and evaluation research literatures provide detailed accounts of the political nature of evaluation, yet both miss vital insights provided by the other. As Mary Daly writes, 'governance is a state-centred perspective and tends to frame change and development in terms of a re-articulation of the state, rather than originating in society' (2003: 125). As a result it can over-state the construction of evaluation as a unitary state-centred practice and miss its contested and complex nature. Contestation can be seen both within the academy and more widely as different 'stakeholders' – in particular, users of services' organisations – compete to use evaluation for different social and political ends. I argue that evaluation should, therefore, be seen as *socially located* and understood as *politically contested*.

The social research literature clearly observes the political struggle over control of evaluation research, but tends to limit its understanding to either the internal *practice* of evaluation or the *post hoc* use of evaluation findings by politicians. In those cases where structural constraints are acknowledged, driven mainly by a critical realism, evaluation theory frequently side-steps the issue of how to position itself within the objective power relations it acknowledges. In contrast, and in line with insights offered by the governance approach, I suggest that evaluation is discursively constructed. Its political meaning is constituted independently of any particular use or practice. However, it is also socially contested and this contestation by different stakeholders takes place within social relations of power.

In order to understand the politics of evaluation, therefore, we need to consider both its role within political discourse *and* its contestation within society. The argument here, then, is not to suggest a particular approach to evaluation is more politically acceptable, but that we must understand evaluation at a number of levels: at the level of formal political strategy; at the level of discursive construction; at the level of informal political processes, especially within those communities which often provide the context for programme evaluation; and at the level of counter-politics associated with attempts to develop radical alternatives.

In the first instance, this article sets out the approach to evaluation from a governance perspective and then goes on to consider the complexities of evaluation as a contested practice and concludes by offering an alternative approach to understanding the 'politics of evaluation', reworking the strengths of these two different traditions. It draws on a range of research projects presented in Taylor and Balloch (2005).

https://doi.org/10.1017/S0047279405009177 Published online by Cambridge University Press

Evaluation as governance

According to Clarke (2003), the central 'problem' of contemporary governance is how an increasingly dispersed state can be 'controlled at a distance'. For Newman (2002), Glendinning *et al.* (2002), Balloch and Taylor (2001), Taylor (2001) and others, contemporary states become more reliant on networks in shaping and delivering public policy and seek new forms of community-based governance. This leads not only to the dispersal but the fragmentation of public services. A resultant hollowed-out state comes to rely on new techniques of government. These are articulated as detached from direct political values and based on the production and consumption of objective evidence of 'what really works' in 'the real world'.

A governance approach helps us understand evaluation as a component of political strategy, even if it is not always openly articulated as such. This strategy is generated by a range of economic and political imperatives: corporate globalisation, the rise of anti-statist ideologies such as neo-liberalism, the fiscal crisis of the state and the fiscalisation of public policy, discourses of 'efficiency' and consumerism, and the public distrust of professional integrity and political ideologies (Clarke, 2003). Governments in this context come to emphasise 'performance management' as a means of controlling quality and standards. Performance 'emerges as something that can be managed and evaluated in the public interest' and 'the performance/evaluation nexus represents a solution to the problems of managing a dispersed and fragmented system at arms length' (Clarke, 2003: 6).

Evaluation is seen as a central component of a resultant new regime of scrutiny, which defrays political responsibility and replaces it with the 'virtuous circle of evaluation, evidence, performance'. This 'turn to evidence' is a new mode of political regulation, and evaluation is part of a wider logic of decentralised control. Evaluation, in this context, relies on an instrumental rationality based on technical evidence. As Newman pointed out when considering the first New Labour term of office, 'policies introduced in the first years of the (New Labour) government . . . emphasised innovation, experimentation and *policy evaluation* to build sustainable long-term change in public services' (2002: 2, emphasis added). Within this new mode of regulation, performance is constantly monitored by an array of state-sponsored new public management techniques.

> The growth of interest in evidence-based policy is linked to a number of developments including the explosion of data, developments in IT, the growth of a well-informed public unwilling to rely on professional judgement and the growth in the size and capability of the research community. These developments have been complemented by the expansion of managerial approaches to the collection and use of *evidence represented in programme evaluation.* (2002: 69, emphasis added)

https://doi.org/10.1017/S0047279405009177 Published online by Cambridge University Press

According to Clarke (2003), the performance culture has become so pervasive that evaluators and evaluatees engage in a deft choreography to produce shared constructions and meanings: they play the game of 'preformativity'. This perspective rightly points us to the construction of evaluation as a strongly self-regulating strategy of governance. It helps move the discussion of evaluation beyond purely methodological considerations and towards the restructuring of public services and the desire of governments to avoid political responsibility for service deficiencies.

A key ideological driver in this shift has been the representation of the public as consumers requiring evidence as a basis for informed 'life-style' service choices. This trend can be traced back in Britain to at least the early 1990s and the consumerism which underscored the reworking of the notion of active citizenship. Since then we have seen the growth of league tables and other devices acting as information sources for rational decision making. As Clarke argues, 'the collection of evaluative and comparative information is understood as making organisations more "transparent" to consumers and making information available that will enable the exercise of rational choice on the part of consumers' (2004: 40). Informed consumer choice then appears a self-evident good. Perri 6, for example, takes such a view as axiomatic when he argues 'few would oppose increased choice for consumers of public services on general principle' (2003: 264). Public interest on this view is articulated around the service user as consumer and the evaluator as evidence bearer. In addition, evaluation and audit are seen as generating improvement and enhancing public accountability. However, as Gubba and Lincoln (1989) observe, a preoccupation with accountability usually means 'someone is about to get blamed'.

The increased focus on accountability also has little to do with real improvement. The 'evaluation, evidence, performance' circle seeks to establish a direct line of accountability between the users of services, constructed as consumers, and welfare services, thus circumventing formal politics. It may be a basis for some changes in relationships between professionals and users of services, but just as consumers enter markets from unequal positions, users of services enter them from social locations, bestowing more or less power and have limited opportunities to use evaluation evidence to influence the service they receive. Users of services have different needs and are afforded differential rights. Their ability to gain information about, access to and shape service delivery is limited by their social location and their power. Against this backdrop, the evaluation of services is not neutral but works in the context of power relations and provides an information service that is formed by interests and power.

This 'meso-narrative' of involvement of the pubic as consumers in debates about the nature of welfare provision is an important element of governmental strategy and is reflected as a statutory duty of consultation in

https://doi.org/10.1017/S0047279405009177 Published online by Cambridge University Press

much 'modernisation' legislation. It aims not only to enhance information and accountability, but the inclusion of users of services is seen as building social capital and developing 'co-governance'. This has been described as a shift from government to governance: from representative to participatory democracy. Participatory democracy is seen as collaborative government and is mirrored by a move towards collaborative or participatory strategies of evaluation research. This shift towards co-governance and participatory evaluation research can be seen as an attempt to align the site of knowledge production with changing strategies of political control. There is a crucial tension, however, in the representative/participatory binary, which is often lost in mainstream political representations of change. Both in politics and in evaluation practice, neither form addresses adequately the question of *presence* and *voice* or – to put it another way – less direct forms of participation. Just as the structures and spaces of participatory democracy may marginalise certain voices, participatory evaluation runs the risk of excluding what it calls 'hard to reach groups' or setting a research agenda dominated by powerful voices.

From a critical governance perspective, Newman points out, within the rationalist model 'there is a clear separation between knowledge production (by experts) and knowledge use (by practitioners)' (2002: 71). There is a third dimension that is missing from this analysis, however: knowledge control by users of services. Evaluation researchers who provide much of the 'evidence base' for this 'knowledge' confront this dilemma which is at the heart of the politics of evaluation – the struggle over expert knowledge and its use – by practitioners, politicians, policy makers and most importantly, users of services themselves.

So where does this leave evaluation from a governance perspective? Certainly it can be seen as a constituent in a new mode of regulation recruited to legitimise new approaches to welfare governance. It is bound up with performance evidence, but seeks to go beyond audit's logic of accountability based on checking against standards towards a focus on improvement. Nevertheless, it is seen as a tool in a state-centred modernisation strategy, driven by national and supra-national political imperatives. Such a perspective allows us to locate evaluation as a regulatory practice within a wider understanding of political change. It also allows us to see its political and social meaning as discursively constructed. This, I believe, is an important insight. There may be a danger, however, that on this view we miss the potential to see the contested nature of evaluation both within the academy and within society. Evaluators come in many political guises and, while it is absolutely the case that evaluation is constrained by the very real economic and political context within which it operates, there are ways in which alternative approaches to evaluation can be developed.

https://doi.org/10.1017/S0047279405009177 Published online by Cambridge University Press

Approaches to evaluation

Debates between social constructionists and scientific realists provide a useful context for understanding different conceptions of politics within the literature on evaluation research. Those who most readily accept the inherently political nature of evaluation – historically, social constructivists – tend to have the most limited conception of the political realm, while those who seek an objective science of evaluation refer most readily to the wider political relations within which evaluation is located. A diversity of ontological and epistemological positions now characterises these theoretical debates, but what is most relevant here is the way in which they have generated different understandings of politics and power.

Much social constructionist evaluation rejects the quest for 'scientific objectivity' because it equates this with the pursuit of a practice independent of political influence, a refusal to acknowledge the political role of the evaluator, and a denial of the contested and relative nature of knowledge. While some extreme versions of scientific realism do subscribe to this position, contemporary realists tend towards a greater understanding of the structural constraints placed upon evaluation by wider political issues and are far from denying the political nature of evaluation.

Early work in Britain, which acknowledged the political dimensions of evaluation, drew on work in the USA with its strong theoretical commitment to pluralism. MacDonald (1974), in developing evaluation studies in education, was one of first in Britain to propose the notion of democratic evaluation. This approach aims to democratise knowledge based on a 'right to know' for all stakeholders. Awareness of educational performance is seen as a citizen right, with the evaluator providing an information service to the community. It is set within a conception of the liberal state, a plurality of interests, and an informed citizenry. Throughout the 1970s and 1980s, the notion of the detached scientific observation of policy and programmes came increasingly under scrutiny with the rise of social constructionism and various forms of post-modernism. By the 1980s, Gubba and Lincoln led the assault on earlier generations of 'scientific' evaluation in the name of social constructionism. They rejected 'the controlling manipulative (experimental) approach that characterizes science and substitute(d) for it a *hermeneutic/dialectic process*' (Gubba and Lincoln, 1989: 44). In their view, evaluation data 'have neither special status nor legitimation, they represent simply another construction to be taken into account in the move toward *consensus*' (1989: 45, emphasis added). They argue that phenomena can only be understood in context and there is no possibility of generalising knowledge from one programme evaluation to other contexts. More recently, Clarke (1999) has argued that evaluators are brokers among different interest groups and stakeholders acting as informants and educators. Clearly, the 'democratic' constructivist approach allows us to understand the situated and

https://doi.org/10.1017/S0047279405009177 Published online by Cambridge University Press

contextual nature of knowledge. It is particularly helpful in understanding the way evaluation functions within the informal political process at the community level and the way in which the voices of politically and socially marginalised or excluded groups can be brought into the policy debate. This approach, however, runs up against the problem of how to achieve consensus if different stakeholder views are all accorded equal legitimacy and when the wider context *beyond* that within which evaluation takes place is 'bracket out'. As evaluators are seen as orchestrators of a negotiation process and facilitators in the achievement of consensus, the political commitment of this type of evaluation is to procedural values such as negotiated agreement.

More recent approaches to 'empowerment evaluation', especially as developed by Fetterman (2001), stress the moral responsibility of evaluators as advocates of the powerless. This 'self-help' approach to evaluation 'is designed to help people help themselves' (Fetterman, 2001: 3). Again the evaluator acts as coach or facilitator for all programme participants and enables collaboration with and among stakeholders. Different stakeholders are brought together to establish improvement goals with the evaluator acting as 'critical friend'. The focus is on self-determination and the aim for participants is the more radical one of 'liberation'. Rather than measuring programme outcomes, evaluators facilitate processes aimed at programme 'improvement'. In this sense there is a deep commitment to community learning from evaluation and of taking control by participants. This approach is avowedly political and recognises the specific context as influencing the nature of the evaluation enterprise. It seeks to enter into that context to build community capacity to take control of the evaluation process. In this context Zimmerman (quoted in Fetterman, 1998: 2) distinguishes between empowerment processes and empowerment outcomes:

> Empowerment processes are ones in which attempts to gain control, obtain needed resources, and critically understand one's social environment are fundamental. The process is empowering if it helps people develop skills so they can become independent problem solvers and decision makers . . . empowering processes for individuals might include organisational or community involvement; empowering processes at the organisation level might include shared leadership and decision making; and empowering processes at the community level might include accessible government, media and other community resources. Empowered outcomes refer to the operationalisation of empowerment.

Cousins and Whitmore, drawing on early work by Foote Whyte, distinguish between two different approaches to contemporary participatory evaluation: what they call 'practical participatory evaluation' and 'transformative participatory evaluation'. Both approaches stress the central role of participant knowledge in evaluation, but transformative participatory evaluation has an explicitly political aim: to enhance participants' understanding of the 'connections among knowledge, power and control' (1998: 8). In this view,

https://doi.org/10.1017/S0047279405009177 Published online by Cambridge University Press

evaluation is seen as a value-driven activity, with a liberatory potential set within a democratic pluralist model of competing interests.

Participatory evaluation with these ideals, however, faces a set of challenges. Cousins and Whitmore talk about the problem of stakeholder diversity and depth of participation. The logic of participation inevitably throws up a diversity of stakeholder interests and there is a danger that this can lead to a relativistic stalemate, where each stakeholder's view is seen as equally plausible. However, if stakeholders are understood as positioned in wider power relations, there is the possibility of employing elements of the 'realist' approach and a more conflict-oriented perspective. In this respect I suggest that evaluators need to approach evaluation with a sociological understanding of the structured bases of inequality in society.

Participation also plays a central role in contemporary discourses around new forms of community-based governance. There is an alignment here between a governance strategy based on consultation and community development, and participatory forms of evaluation seeking to develop community learning and empowerment. This conjuncture may provide *a space to pursue radical approaches to evaluation*. However, Newman cautions that, rather than delivering genuine change, there is a danger that the importance attached to participation is really more about legitimisation of new forms of governance 'than about a genuine willingness to transform decision-making processes' (2002: 137). 'On the one hand, participation processes (open) up new spaces which could be captured by user groups, voluntary organisations and community groups, claiming a stronger role in decision-making' (2002: 138), but at the same time there is a danger that the basis of participation may itself be exclusionary. This is especially so when participation is undertaken on what are seen as stigmatising terms: that is, the acceptance by participants of a disempowered identity or social location. Newman cites an example of programme participant/service users who define themselves as lacking money, but who are reluctant to see themselves as socially excluded. Their participation in an evaluation project may validate an outcome based around a focus on exclusion/inclusion and a subordinate status. Participatory evaluation, therefore, needs to avoid the danger of assigning some totalising identity (Taylor, 1998) – in this case 'socially excluded' – which further disempowers those in need. Involving stakeholders in this context, especially disempowered stakeholders, means listening to their self-conceptions and the meanings they attach to social programmes and interventions in their lives.

Community empowerment in this context is a complex issue. Debates about community safety, for example, illustrate how some programmes may lead to a positioning of outsiders or indeed marginalised insiders as a threat to 'community values'. Restricted access to public spaces, gated communities, anti-social behaviour orders and neighbourhood watch are examples of 'community empowerment' which can be just as exclusive as inclusive. Evaluators of such

https://doi.org/10.1017/S0047279405009177 Published online by Cambridge University Press

projects would need to consider very carefully just *which* members of the community should be involved and how certain stakeholder views might frame the research. These are awkward questions for evaluators wishing to enable participation and ones that need to be set within a wider conception of social relations of power.

Newman warns that in certain circumstances:

> participation may worsen relationships between users and providers, between communities and public bodies, between citizens and government . . . communicating clearly about the aim and scope of consultation, and giving feedback about the outcomes, are clearly important. But the political tensions in the process, and the potential conflict these give rise to, cannot be massaged away: more effective management cannot solve problems in the political domain. (2002: 139)

One might paraphrase and say that effective strategies for participatory evaluation cannot always change the fundamental situation in which users of services find themselves and may even reinforce negative stereotypes. As Clarke in his review of evaluation research suggests, empowerment evaluation may 'give a false impression of empowerment' (1999: 29). Cousins and Whitmore (1998), in discussing what they call the 'depth' of participation, state there are important considerations about representativeness of stakeholders to be borne in mind. Many social programmes and their evaluations can be captured by the established activists or those whose voice is dominant: often the most organised groups in the community. The question then is on what terms will users participate in evaluations? Should evaluators work with a broader notion of participation than is found in more traditional understandings of representative or even participatory democracy, and how should evaluators understand their own relationship to their research practice?

Fetterman and other empowerment evaluators see the evaluator as facilitator or negotiator. Indeed, Fetterman goes even further and urges evaluators to campaign on behalf of the powerless. However, we immediately confront the fact that evaluators are faced with value choices, which is not an easy position for some constructivist evaluation. A second problem also confronts the relativist version of constructivism. If all knowledge is particular and narrowly contextual, there is no hope of making any generalisations about the usefulness of policies or programmes from evaluation studies or generalising about the way social programmes can empower or disempower particular groups within *society*. This critique also divides those who take a scientific realist approach to evaluation research.

Writers such as Martin Hammersley caution against arguments that research should pursue political ends: 'to the extent that they become widely accepted, they negate social researchers' attempts to preserve some autonomy from the state and from powerful social interests, thereby destroying the conditions in which

https://doi.org/10.1017/S0047279405009177 Published online by Cambridge University Press

research can flourish, perhaps even threatening its survival' (1995: ix). While this may be useful in guarding against incorporation into 'state-centred modes of regulation', Hammersley is most dismissive of those positions most likely to oppose state incorporation – specifically 'feminist, anti-racist and postmodernist approaches' – for their value commitment. Instead, he locates independence in the political autonomy of the researcher. This is perhaps an example of what Simons (1987) calls the (scientific) 'autocratic' approach to research and evaluation: a model based on sustaining a detached scientific legitimacy. However, Pawson and Tilley offer a rather more subtle realist approach to understanding evaluation. They certainly subscribe to the notion of the independence of judgement in research, but these judgements must be about power and politics. The context within which evaluation takes place 'requires the researcher to generate some means of making *independent* judgements about the institutional structure and power relations within a program' (1997: 23). This, they argue, is a position which 'constructivism cannot contemplate'. They provide a useful critique of the limits of contextual relativism and propose a *philosophical* justification for seeking generalisations. However, it remains rather abstract. Their reference to 'some means' of making independent judgement is illuminating. While they go beyond the limits of pluralism and hint at a 'politics of evaluation' in comments such as 'the asymmetries of power . . . allow some people to advance their ideas while others have choices foreclosed' (1997: 23), they do not spell out the means by which the evaluator evaluates these power relations or, to put it another way, adopts an explicit value position.

Evaluation for whom?

In an overview of realistic evaluation, Tilley asks the question 'what works, for whom in what circumstances' (2000: 4). But who should be asked and about what? These are as much political as methodological issues. The questions and the answers they generate are likely to be influenced by an understanding of the purpose of social programmes and the values underlying them. Jennifer Greene asks the question of evaluation, 'what do meaningful representations of program quality contribute to the larger conversation about the public purposes of social investments?' (1999: 161), thus explicitly raising the issue of political values. Evaluators have some opportunities for saying 'how' but are generally more limited in their options for saying 'what' is to be evaluated and 'why'. But the question, 'who is evaluation for?' needs to be addressed.

There is an established understanding within the literature of the problem of stakeholder conflict and the relative power between them – including funders, service providers, and users of services or other programme 'beneficiaries' – and some attention is given to who sets the agenda. This issue is often characterised as 'top–down' evaluation versus 'bottom–up' evaluation. The purposes of

https://doi.org/10.1017/S0047279405009177 Published online by Cambridge University Press

'top–down' evaluation are usually to do with programme legitimation and tend to be outcome oriented. 'Bottom–up' evaluation is said to be driven by the needs of the users. However, whether evaluation is inspired by the 'top' or the 'bottom', some stakeholders may have more at stake than others, and the balance of interests may shift as evaluations unfold. As Themessl-Huber and Grutsch write, 'some stakeholders . . . will have premonitions about resulting job losses and devastating budget cuts' (2003: 93). We might add, however, some might have concerns, as described by Newman above, that their involvement may further stigmatise or simply act as a legitimating device for advocates of consultative decision making and inclusive policy making, or that they may have a range of interests in a programme, which cannot simply be captured by an assigned social identity and its assumed attendant 'interests'.

Beresford (2002, 2005) proposes a radical approach to this issue. He argues for 'user controlled' research and evaluation. Much participatory evaluation is seen as limited in its conception of involvement and based on the consumerist view of the user. Against this, Beresford – drawing on work by Oliver and others in relation to the disability movement – poses a view of research which is controlled by users of services themselves:

> a number of differences can be identified between these two approaches to participation. The first approach generally starts with policy and the *service system*; the second is rooted in people's lives. Both approaches may be concerned with bringing about change and influencing what happens. However in the consumerist approach, the search is for external input which the initiating agencies (state, service providers, or policy makers) *themselves* decide what to do with. The democratic approach is concerned with ensuring that participants have the direct capacity and opportunity to make change. This latter approach highlights issues of power and the (re)distribution of power. These are not explicit concerns of the consumerist model of involvement. (Beresford, 2002: 97)

In this view, users take control of evaluations supported by researchers. In more traditional forms of participation, there is a danger that the evaluation research is simply concerned with regulating people's lives. Hogwood and Gunn exemplify this approach when they argue: 'it may not be appropriate to ask the target group for their particularistic perspective, especially when they are the target of regulatory policies' (1984: 223). They appear to have no qualms about policy evaluators complying with this strategy of regulation. Beresford, on the other hand, argues 'research approaches from service users tend to be framed in terms of emancipation and liberation rather than participation, with participation seen as a necessary but not sufficient condition for liberatory or emancipatory research' (2005: 78). The traditional model is linked directly to changing forms of governance.

> the managerialist/consumerist approach, developed in the last quarter of the twentieth century out of the commitment of the emergent political New Right to more market-driven approaches in both political ideology and public policy, with a devaluing of state intervention and provision.

https://doi.org/10.1017/S0047279405009177 Published online by Cambridge University Press

> It placed a new emphasis on purchase of services and individualised arrangements and responsibility for health and welfare. The recipient of welfare was reframed as a consumer (albeit sometimes an involuntary one) and a rhetoric of individual 'choice, voice and exit' developed. With the emphasis on improving the 'three Es': economy, efficiency and effectiveness, user involvement was framed in market research terms of 'improving the product' through market testing and customer feedback. (p. 79)

For user-controlled research and evaluation, 'emancipation is the goal, not necessarily the achievement, of such research' (Beresford, 2002: 100). With its origins in exactly the sort of radical political movements which Hammersley rejects as a value base for research, this approach tends to be tied to a wider political project. There is a marked difference between this and earlier forms of participatory evaluation. The latter tend to look at the specific contextual political issues associated with any evaluation project and consider how participants can be included within it. User-controlled evaluation focuses on wider rights-based issues with social welfare movements determining their own research and evaluation agenda. The contrast is not simply an evolutionary difference in which more recent forms of user-controlled evaluation are more political or more radical. These approaches involve different political strategies. Empowerment evaluation tends to focus on variegated communities with a range of stakeholders who may or may not see themselves as service users with clearly articulated needs and are quite frequently disengaged from formal politics. Beresford's approach, on the other hand, points us to the broad political nature of evaluation and seeks to locate an alternative politics of evaluation in the rights and needs of disempowered but politically active social groups. This approach may lend itself to rights-based political campaigns more readily than to specific locality-based programme evaluation. It may not be so applicable where there is no readily identifiable single user group, or where there are several competing disempowered groups within a programme, or where users of services do not wish to identify so explicitly with a particular social identity or become part of a social movement. In these contexts radical evaluators may need to adopt insights from the 'community development' tradition, with its emphasis on capacity building among a range of more diffuse community constituencies alongside the more universal rights-based campaigns described by Beresford.

While there is a contested politics at the heart of evaluation, then, I argue that this is more complex than either the governance or the evaluation research literature allows. On the one hand, despite the very real attempts to construct evaluation as a state-centred mode of regulation, there is a counter politics at work in the approach of radical service users' movements and emerging alternative approaches within evaluation practice. However, on the other, user-controlled research which offers a strong counter to the state-centred perspective may be most relevant to universal rights-based campaigns by politically active groups. The more fragmented and dispersed groups of service users who are recipients of

https://doi.org/10.1017/S0047279405009177 Published online by Cambridge University Press

locally based programmes may need a sensitive set of evaluation practices, which can be adapted and moulded to particular circumstances.

A radical community-based approach to evaluation may have strengths in asking in what ways does the design of the evaluation accommodate different interests and allow for different voices to be heard? There can be a danger in this approach, though, of *essentialising* the interests of stakeholder groups. Voice, presence or other forms of participation and representation within communities are complex processes, which may be characterised by fluidity as much as constancy of interest. Gail Lewis, in a discussion of difference and diversity, has suggested that difference is inseparable from issues of power and inequality. 'As a consequence, the multiple forms of social difference in relation to class ethnicity, gender, age (dis)ability, religion and sexuality become inflected through dynamics of power' (2003: 92). Diversity, in this context, becomes something to be achieved as different voices struggle for presence. In this struggle, however, interests develop and change as they are mediated by a range of different experiences and social locations. Difference is not, therefore, naturalistic but part of an ongoing process of development and realisation. This helps us understand that for evaluation it is not simply a matter of obtaining the pre-set views of established groups or individuals but of being sensitive to the multiplicity of positions and subjectivity and to the role of the professional evaluator in supporting different groups and individuals in identifying their needs. In this context we need to ask whether the evaluation design accommodates changes along the way as power dynamics shift and in what specific ways the design allows different forms of participation? Themessl-Huber and Grutsch argue, for example, that for participatory evaluation, 'conscious and active participation of stakeholders is not a necessary pre-requisite for empowering and emancipating the participants . . . stakeholders are not required to have a vital interest in the evaluation. They learn and acquire knowledge by being informed about the progress of the project' (2003: 94). Careful consideration therefore needs to be given to differential strategies for involving different stakeholders and the relative strength and depth of their involvement.

The dangers of attributing fixed interests and focusing on particular groups as research subjects was illustrated by John H. Stanfield in his 1999 address to the American Evaluation Society. He launched a devastating critique of the racialisation of evaluation within the American Academy (Stanfield, 1999). He pointed out that the way evaluation had evolved predisposed researchers not only to ask negative questions about black experiences but *not* to ask questions about the strengths of communities, which might shed a more positive light on black people (Stanfield, 1999: 421). Evaluation studies focused on the negative aspects of black people's experiences casting them in passive victim mode. He saw the racialised structures of white-dominated research as reinforcing negative stereotypes of poor black people. Ann Oakley has made a similar point in relation

https://doi.org/10.1017/S0047279405009177 Published online by Cambridge University Press

to the way in which research is gendered: 'neither methods nor methodology can be understood except in the context of gendered social relations' (Oakley, 2000: 4).

Avoiding closed paradigms of evaluation in which participants are constructed according to pre-established conventions or are subject to negative totalising identities is both a political and methodological challenge for policy evaluation. A range of strategies has been developed within the evaluation community, which seek to address these issues, especially in writings that span action research and evaluation: MacLure and Cronach's (1997) 'report and respond' approach; the developing tradition of dialogical evaluation associated with a range of writers predominantly from the USA (see Abma, 2001, for an overview); Deepwell and Cousin's (2002) conversational approach; Cousin *et al.*'s (2005) use of 'appreciative enquiry' offer a variety of ways forward. In 'appreciative enquiry', evaluators start from asking positive rather than negative questions about participants' experiences and seek to build a strong case of 'what works' for people on the ground, and as such offer some possibility of this approach gaining legitimacy within formal political circles. Cousin *et al.* (2005) give examples of how, as evaluators, they invited participants to prepare stories of their experiences citing moments when 'things clicked' or really worked, rather than focusing on the negative and reinforcing the problem-centred focus of disadvantage. Clearly, oppressive practices and disempowerment cannot be ignored, but, if these are situated within a wider understanding of social inequality and social relations which sets the context, then evaluation practice can focus more on identifying the positive needs of users of services.

The process by which diverse and marginalised groups can find a voice has also been considered recently by Graham and Harris (2005) in their approach to a participatory evaluation of a New Deal for Communities (NDC) programme. In their evaluation project, 'local people are integrally involved in it, it builds up capacity within that community and is based on the principles of empowerment, equity and democracy: all of which are guiding principles in the whole programme. Additionally, it affords local people the opportunity to learn new skills and encourages them to engage with local issues' (p. 100). The most popular and widely accepted research methods – surveys, interviews, focus groups and so on – need to be set to one side, they argue, in favour of more creative approaches. 'These are likely to be far less exclusionary for some of the more readily quashed voices (for example, people with English as an additional language, people with learning difficulties or disabilities, children and young people) who can communicate in ways in which they are perhaps more comfortable' (p. 101). They use a variety of methods to elicit voice, including photography, sculpture, textiles, storytelling, social mapping and drama.

As Graham and Harris argue, creative approaches can offer a number of advantages for participatory evaluation. As well as being less inhibiting than

https://doi.org/10.1017/S0047279405009177 Published online by Cambridge University Press

some traditional methods, they allow the creator to express emotions, experiences or aspirations (*ibid.*). As Paul Hoggett has argued, much contemporary social thought fails to acknowledge the importance of the emotional life. Instead of an 'over-socialised' conception of 'man' characteristic of much recent social thought, Hoggett proposes a 'generative' approach to understanding welfare in which individual and social development 'is seen as the realisation of our aesthetic, moral, spiritual and practical capacities' (2000: 2). His approach adds a subjective dimension to struggles over redistribution and social justice. It includes consideration on the quality of social relations between men and women, between generations, between and within families and communities and between users of welfare services and those who design and deliver them' (2000: 208). It is about what Titmuss described as the texture of social relationships. From such a perspective, evaluation can be approached as a means of allowing users of services the space for the expression of feeling and for making sense of their experience. The use of creative evaluation methods which are highly symbolic may be helpful in such an inter-subjective process. As Graham and Harris suggest, 'people often feel less inhibited expressing themselves through a medium other than speech or writing . . . It is visually impactive, so participants can more readily take part in the interpretation of the pieces – both theirs and those created by others' (2005: 103). And, while dealing with emotional responses, 'they offer safer spaces for expression by depersonalising emotional stories (for instance, when working with metaphor)' (ibid.). They conclude, however, that, although creative approaches appear to offer great promise, their application in evaluation remains in its infancy. It should also be remembered that unless they are introduced and managed well, they could potentially be exclusionary to some groups (ibid.). Finally, we also need to be aware that finding suitable processes and spaces within which voices can be expressed is only part of the issue. Voices also need to be heard by those in power. There are many institutional and professional barriers which can mean that expressed views flounder on the rocks of established practices or the silence of the powerful. A critical evaluation, therefore, needs to address wider power relations and how the voices of marginalised groups can illicit a response which is more than tokenistic.

There are, indeed, many more constraints on the development of a genuinely empowering approach to evaluation. These include the time required of participants to engage in programme evaluation, the possible cost both to participants and to the project (genuinely inclusive and creative approaches tend to be resource intensive), motivation and confidence. Additionally, it may be difficult to design genuinely empowering approaches to programme evaluation, if the programme or policy itself has been designed and delivered in an exclusionary manner. However, these complex issues are being addressed by evaluation researchers and advocacy movements committed to a wider understanding of the politics of evaluation.

https://doi.org/10.1017/S0047279405009177 Published online by Cambridge University Press

Conclusion

Understanding of evaluation is increased if it is set within the context of changing political imperatives. I believe that recent governance theory sets evaluation in just such a context and helps us move beyond pure methodological considerations of evaluation practice to a consideration of discursive construction of evaluation as a constituent of a new mode of regulation. However, if there is an over-reliance on a state-centred perspective, the politically contested nature of evaluation can be lost and struggles by users of services can be overlooked. Traditional approaches to participatory evaluation tend to rely on a democratic pluralist model of politics and fail to consider both the wider context within which particular political struggles take place and assume that participants can be included in research designs and agendas set by funders and/or professional evaluators. A radical tradition, however, has emerged in which users of services seek to set the agenda and control evaluation. This has developed out of rights-based user movements and provides a way of locating users' experiences at the centre of evaluation. However, such an approach itself needs to be sensitive to the complexity of user experience and the need to develop research methods that allow for uncertainty, inter-subjectivity and complexity. Service user movements and radical evaluators have together begun to develop a range of creative methods which may offer a way forward for an approach to policy and programme evaluation which is genuinely empowering and based on the needs of those for whom policy should be developed.

References

6, P. (2003),'Giving consumers of British public services more choice: what can be learned from recent history?', *Journal of Social Policy*, 32: 2, 239–70.

Abma, T. (2001), 'Opening thoughts: special edition on dialogical evaluation', *Evaluation*, 7: 2, 155–63, London: Sage.

Archer, M. (1995), *Realist Social Theory: The Morphogenic Approach*, Cambridge: Cambridge University Press.

Balloch, S. and Taylor, M. (2001), *Partnership Working: Policy and Practice*, Bristol: Policy Press.

Beresford, P. (2002), 'User involvement in research and evaluation: liberation or regulation?', *Social Policy and Society*, 1: 2, 95–106.

Beresford, P. (2005), 'Service user involvement in research and evaluation: issues, dilemmas and destinations', in S. Balloch and D. Taylor (eds.), *The Politics of Evaluation*, Bristol: Policy Press.

Bhaskar, R. (1975), *A Realist Theory of Science*, Brighton: Harvester.

Clarke, A. (1999), *Evaluation Research*, London: Sage.

Clarke, J. (2003), 'Performing for the public: evaluation, evidence and evangelism in social policy', Paper presented to the Social Policy Association annual conference, Middlesborough, July.

Clarke, J. (2004), 'Dissolving the public realm? The logics and limits of neo-liberalism', *Journal of Social Policy*, 33: 1, 27–48.

Clarke, J., Gerwitz, S. and McLaughlin, E. (eds.) (2000), *New Managerialism, New Welfare?*, London: Sage.

https://doi.org/10.1017/S0047279405009177 Published online by Cambridge University Press

Cousin, G., Cousin, J. and Deepwell, F. (2005), 'Discovery through dialogue and appreciative inquiry: a participative evaluation framework for project development', in D. Taylor and S. Balloch (eds.), *The Politics of Evaluation*, Bristol: Policy Press.

Cousins, J. B. and Whitmore, E. (1998), 'Framing participatory evaluation', in E. Whitmore (ed.), *Understanding and Practicing Participatory Evaluation: New Directions for Evaluation*, San Francisco: Jossey-Bass.

Craib, I. (1997), 'Social constructionism as a social psychosis', *Sociology*, 31: 1–15.

Daly, M. (2003), 'Governance and social policy', *Journal of Social Policy*, 32: 1, 113–28.

Deepwell, F. and Cousin, G. (2002), 'A developmental framework for evaluating institutional change', *Educational Developments*, 2: 5.

Fetterman, D. (1998), 'Empowerment evaluation and the Internet: a synergistic relationship', Current Issues in Education (On-line) http://cie.ed.asu.edu/volume1/number4.

Fetterman, D. (2001), *Foundations of Empowerment Evaluation*, California: Sage.

Glendinning, C., Powell, M. and Rummery, K. (2002), *Partnerships, New Labour and the Governance of Welfare*, Bristol: Policy Press.

Graham, K. and Harris, A. (2005), 'New deal for communities as a participatory public policy – the challenges for evaluation', in D. Taylor and S. Balloch (eds.), *The Politics of Evaluation*, Bristol: Policy Press.

Greene, J. (1999), 'The inequality of performance measures', *Evaluation*, 5: 2, 160–72.

Gubba, E. G. and Lincoln, Y. (1989), *Fourth Generation Evaluation*, California: Sage, p. 294.

Hammersley, M. (1995), *The Politics of Social Research*, London: Sage.

Hoggett, P. (2000), *Emotional Life and the Politics of Welfare*, Basingstoke: Macmillan.

Hogwood, B. and Gunn, L. (1984), *Policy Analysis for the Real World*, Oxford: Oxford University Press.

Lewis, G. (2003), Difference and social policy', in N. Ellison and C. Pierson (eds.), *Developments in British Social Policy 2*, Basingstoke: Palgrave Macmillan.

MacDonald, B. (1974), 'Evaluation and the control of education', in B. MacDonald and R. Walker (eds.), SAFARI 1 Innovation, Evaluation, Research and the Problem of Control, University of East Anglia, Norwich.

MacLure, J. and Cronach, A. (1997), *Educational Research Undone: The Postmodern Embrace*, Buckingham: Open University Press.

Newman, J. (2002), *Modernising Government: New Labour, Policy and Society*, London: Sage.

Oakley, A. (2000), *Experiments in Knowing: Gender and Method in the Social Sciences*, Cambridge: Polity Press.

Pawson, R. and Tilley, N. (1997), *Realistic Evaluation*, London: Sage.

Schilling, C. (1997), 'The undersocialised conception of the embodied agent in modern sociology', *Sociology*, 31: 737–54.

Simons, H. (1987), *Getting to Know Schools in a Democracy*, London: The Falmer Press.

Stanfield, J. H. (1999), 'Slipping through the front door: relevant social scientific evaluation in the people of colour century', *American Journal of Evaluation*, 20: 3, 415–31.

Taylor, D. (1992), 'A big idea for the nineties? The rise of citizen's charters', *Critical Social Policy*, 33: 87–94.

Taylor, D. (1998), 'Social identity and social policy; engagements with postmodern theory', *Journal of Social Policy*, 27: 3, 329–50.

Taylor, D. (2001), 'Breaking down barriers: reviewing partnership practice', Health and Social Policy Research Centre, University of Brighton, Brighton.

Taylor, D. and Balloch, S. (2005), *The Politics of Evaluation*, Bristol: Policy Press.

Themessl-Huber, M. and Grutsch, M. (2003), 'The shifting locus of control in participatory evaluations', *Evaluation*, 9: 1, 92–111.

Tilley, N. (2000), 'Realistic evaluation: an overview', paper presented at the founding conference of the Danish Evaluation Society.

Zimmerman, M. A. (2000), 'Empowerment theory: psychological, organizational, and community levels of analysis', in J. Rappaport and E. Seidman (eds.), *Handbook of Community Psychology*, New York: Kluwer.

https://doi.org/10.1017/S0047279405009177 Published online by Cambridge University Press