"We know it works…": The Troubled Families Programme and the Pre-Determined Boundary Judgements of Decontextualised Policy Evaluation

Abstract

This article draws on the Troubled Families Programme (TFP) to highlight the ways in which particular contexts – such as socioeconomic and symbolic structures – are neglected in forms of evaluation with an establishment orientation. The article problematises two key aspects of decontextualised evaluation: firstly, the privileging of pre-determined relations of cause and effect; and secondly, the unproblematized framing of policy problems. More contextualised forms of evaluation are presented as a way to open up boundaries of investigation. Lastly, it is argued that an anti-naturalist foundation for evaluation can broaden the scope of learning beyond the original framing of a policy.

Key Words: policy; evaluation; troubled families programme; anti-naturalist; boundary critique

Introduction

Policy evaluation draws on different social research methods to systematically investigate the design, implementation, and effectiveness of an intervention (Rossi, Freeman and Lipsey, 1999, p.20). It can provide the means through which to develop an evidence base to ensure that social policies are effectively delivering what they intended to do (Sanderson, 2002). The key to effective evaluation is seen as being able to assess the impact of a policy intervention and to identify the cause of any changes. Evaluation has the potential to generate critical knowledge to improve social policy in the future.

As most critical theorists of evaluation have pointed out, evaluation is not value-free (Weiss, 1979; Guba and Lincoln, 1989; Taylor and Balloch, 2005; Schwandt, 2007; Mertens and Wilson, 2012). Evaluation studies are shaped through existing political, social, and scientific discourses and practices, which shape the types of evidence that are seen as most robust (Taylor, 2006: 248). Weiss argued that political influences mean that policy evaluation predominantly has an 'establishment orientation' (1993). Weiss (1993: 101) identifies three main political factors that influence an evaluator: firstly, evaluation is most often commissioned by the agency that is responsible for the programme - rather than the people who are the recipients; secondly, the evaluator interprets data in light of the politics that have shaped how the policy has been originally framed; and thirdly, that the primary audience for the evaluation report are policy-makers and not the recipients or the public.

Establishment-oriented evaluations evidence outcome judgements (either from a quantitative or qualitative process) on an intervention, rather than considering a value-critique of the aims of the policy (Taylor, 2006). The values that are internal and associated with the intervention are taken as a given through establishment evaluation. The external values that shape the construction of policy 'problems' and 'solutions' remain unquestioned and outside the scope of evaluation. For Stake (2000: 103), evaluators should 'attend first to their contracts and professional obligations'. There is little scope to question the normative foundations of policy, which results in an uncritical adoption of the boundaries of investigation for an evaluation.

This article problematises the decontextualised approach practised through an establishment-oriented evaluation. A decontextualised approach to evaluation is based on an understanding of participants as objects rather than subjective human agents, locating them outside of a particular context - whether that is socioeconomic structures, symbolic structures or the context of place and time (Krumer-Nevo and Sidi, 2012). The UK Troubled Families Programme (TFP) is used to illustrate the critique of a decontextualised approach to evaluation. The article explains how a decontextualised approach to evaluation is rooted in a naturalist philosophy of science and demonstrates how neglecting contextual factors can limit the scope for learning. Finally, we argue that more contextualised evaluation can broaden the boundaries of investigation and disturb how policy 'problems' have been constructed.

The Troubled Families Programme

'We know it works because we've already looked at studies that show that this works, basically, and also I've met countless families that have been turned around.' (Casey, 2013)

The TFP was established by the Conservative-Liberal Democrat coalition government in December 2011, in response to the riots that occurred in parts of England in August of that year. Even while the riots were ongoing, politicians and the media engaged in an 'orgy of family-blaming' (Gillies and Edwards, 2012: 432), ignoring wider determinants such as poverty, cuts to local services, and a long history of black men dying at the hands of the Metropolitan Police. The programme set out to 'turn around' the lives of the 120,000 most 'troubled families' (characterised by crime, anti-social behaviour, truancy or school exclusion and 'worklessness') through a 'family intervention' approach which advocates a 'persistent, assertive and challenging' way of working with family members to change their behaviours but, crucially, not their material circumstances. The programme is now in its second phase, working with more families identified using additional criteria, and is set to work

with over half a million families in total by 2020. The expansion of the programme was announced just one year into the first phase (DCLG, 2013) and before any evaluation had been carried out.

The programme operates using a Payment by Results (PbR) model, which rewards local authorities when they begin work with a 'troubled family' and then provides further funding when 'significant and sustained progress' that is measured against pre-determined criteria has been achieved, or when an adult moves off out-of-work benefits and into 'continuous employment'. Loopholes in the PbR model in the first phase of the programme meant that local authorities claimed that many families who managed to 'turn around' their lives themselves, without any input or support from family workers on the programme, were direct and quantifiable successes of the TFP (Crossley, 2018).

The evaluation of the first phase of the programme was commissioned by the government shortly after the programme began. The consortium chosen to carry out the evaluation included a number of researchers and organisations that had carried out evaluation of previous Family Intervention Projects (FIPs) and was led by Ecorys - a private sector company that 'provide custom designed research and evaluation solutions' (Ecorys, no date). When the evaluation was announced, the then Communities Secretary Eric Pickles stated:

> We know that the Troubled Families programme is already transforming the lives of thousands and improving the communities around them by tackling truancy, youth crime, anti-social behaviour and worklessness, as well as reducing costs to the taxpayer. However it is important we learn the lessons of this work for the future and leave a legacy beyond the lifetime of this programme in 2015. This study will help do that by looking at what works most effectively with troubled families and how we best spend public money on turning them around (DCLG, 2013).

In August 2016, there were reports that the evaluation had been 'suppressed' (Cook, 2016) because it did not show the impact that the government were hoping it would. The government had claimed that the first phase of the programme had been 99 per cent successful in 'turning around' the lives of nearly 120,000 'troubled families'. In October 2016, just two days before a parliamentary inquiry into the TFP began, the evaluation was published.

There were a number of different streams to the evaluation and a number of different reports published, including: a report drawing on data from a family survey; a report using 'family monitoring data' collected by local authorities; a report on families' experiences and outcomes; a report on the 'process' of implementing the TFP - drawing on interviews and discussions with local authority staff; and a report on the impact of the programme. Much of the press coverage of the evaluation focused on a key paragraph from the national impact study:

> The key finding is that across a wide range of outcomes, covering the key headline objectives of the programme - employment, benefit receipt, school attendance, safeguarding and child welfare - we were unable to find consistent evidence that the Troubled Families programme had any significant or systematic impact. That is to say, our analysis found no impact on these outcomes attributable to the programme. The vast majority of impact estimates were statistically insignificant, with a very small number of positive or negative results. These results are consistent with those found by the separate and independent impact analysis using survey data, which also found no significant or systemic impact on outcomes related to employment, job seeking, school attendance, or anti-social behaviour. This gives us further confidence in the reliability of our results (Bewley et al, 2016: 20).

Other elements of the evaluation received much less press attention, despite them also highlighting shortcomings of the programme and a duplicitous depiction of the families entering it: the process evaluation report highlighted the extent to which the TFP was driving 'service transformation' within local authorities, with the results being ambiguous at best; and the Family Monitoring Data report showed that the clear majority of 'troubled families' that local authorities worked with were not actually that troublesome or anti-social, despite the government rhetoric surrounding them.

The Family Survey Data report (Purdon and Bryson, 2016) provided support to the National Impact study findings that no impact could be attributable to the programme. The report, based on responses from 495 families who had been on the programme for 9 months and a comparison group of 314 families who had just started on the programme, was unable to find any impact attributable from the programme, based on responses from families:

> We found very little evidence that the Troubled Families Programme significantly affected the outcomes of families around nine months after starting the programme. The statistically

> significant improvements we did identify relate to the perceptions of main carer respondents in the Troubled Families group about how they were coping financially, and more generally about how they felt they were faring, and their expectations for the future. There were no positive (or negative) impacts identified for housing, employment and jobseeking, anti-social behaviour and crime, school behaviour and attendance, health, drug or alcohol use, family dynamics or well-being (Purdon and Bryson, 2016: 24).

The evaluation of the second phase of the programme is being carried out by the private sector consultancy firm Ipsos MORI, and the Office for National Statistics (ONS). The first sets of research findings and annual reports relating to the second phase of the programme chose, in lieu of any positive results, to focus on the role the programme was playing on 'service transformation' (DCLG, 2017) and promised a new emphasis on addressing 'worklessness' (DWP, 2017). The most recent findings from the evaluation were published in March 2019 . The evaluators state that 'overall … the programme is having no or limited impact on individuals claiming out-of-work benefits' and there is 'no difference between the programme and comparison group for employment outcomes'.

There are many issues with the TFP evaluation and the programme itself (Crossley, 2018; Hargreaves et al, 2019), but the focus for this article is on the ways in which the TFP evaluation was rooted in a decontextualised approach, set within boundaries set to mirror the original framing of the policy. The TFP evaluations almost completely neglect wider issues of poverty, austerity and welfare reform, and some of their potential consequences. The impact of changes to benefits entitlements, or increased conditionality, were not covered in any of the 765 pages of the seven reports of the first phase evaluation. The phrase 'welfare reform' does not feature at all and 'poverty' appears just twice. Poverty is not mentioned at all in the third tranche of reports (580 pages), despite 66% of families reporting, via the family survey, that their household income was below £12,500 - well below the poverty line for families with children (MHCLG, 2019b: 16). Austerity, mentioned in just two of the first phase reports, was presented as an issue that 'helped to sharpen the focus about the value of family intervention' and 'provided the additional impetus for professionals adopting whole family working' (Day et al, 2016: 24), but it was not considered as an issue that might have had an impact on families themselves. Discussions of poor and precarious labour market conditions, cuts to local authority services for children, young people and families, and inadequate housing provision are almost completely neglected in the reports. The following section identifies the roots of a decontextualised approach to evaluation in naturalist philosophies of social science research.

Naturalist foundations of decontextualised evaluation

Evaluation with an establishment orientation is rooted in a naturalist philosophy of social science. Bevir and Blakely (2017) identify a 'philosophical conflict' between naturalist and anti-naturalist philosophies of social science research. Naturalism arises from a belief that the natural and social worlds are similar enough that they should be studied in the same way (Bevir and Kedar, 2008). Naturalist explanations of the social world are 'causal, ahistorical, and invariant' (Bevir and Blakely, 2016: 33). The constitutive role of meanings in the social world and particular contexts are not investigated through a naturalist approach to evaluation. As a result, issues such as socioeconomic structures, symbolic structures or the situation of place and time are considered to be outside the boundaries within which evaluative judgements will be made. In contrast, anti-naturalism is the belief that human agency makes social inquiry incompatible with the approach of natural science; narratives can be contested and context should be considered as an important dimension of social research (Bevir and Blakely, 2016: 31).

Naturalist social inquiry is concerned with 'discovering' Humean causation to create a predictive science of society. Establishment evaluations that are based on probabilistic statistical procedures aim to evidence linear relationships of cause and effect between the programme and anticipated outcomes. This is based on what is known as the Humean regularity conception - 'the reason that we know that one event has caused another is because the first event took place before the other – regularity of succession' (House, 2001: 311). The aim of the regularity conception is to find a linear relation of cause and effect, using statistical procedures, to evidence that activities A, B, and C will produce the results of X, Y, and Z. If the relationship between A, B, C (the intervention activities) and X, Y, Z (the outcomes) happens regularly enough, then the intervention can be said to have 'worked'. The regularity conception of causation is the basis for testing how an intervention has 'solved' a defined policy problem.

The construction of 'variables' is a foundation of naturalistic inquiry. Evaluations based on a naturalist philosophy of social science use the relationships between variables to predict outcomes and evidence effectiveness. Activities that take place through an intervention are reduced to inputs, and an assessment is made about how regularly these inputs have led to changes in outcome variables (Porter, 2015). To be able to 'discover' causal relationships, naturalist social scientists atomise aspects of society into discrete units of analysis, which can then be correlated and

measured. Measurement of these variables is intended to be systematic and repeatable (Bevir and Kedar, 2008: 511).

A naturalist construction of variables requires an atomisation of the social world in ways that are problematic as it abstracts meaning, agency, and contingency. Atomistic reification occurs to turn intervention inputs into 'objective' independent variables (Bevir and Blakely, 2016: 35). These variables are separated from the participants' interpretations of an intervention and are therefore reified as 'free-floating' and value-free 'facts' (Bevir and Blakely, 2016: 37). The original framing of the policy is thus depoliticised. Another form of abstraction known as essentialism occurs when evaluators create a series of core outcome variables, which can be ascribed commonly across research participants. Evaluators can then decide to exclude, or include, cases on the basis of whether or not they exhibit the determined core attributes. A weaker form of essentialism includes more scope to assess the degrees to which participants fit with the core attributes, but the same fundamental problem exists (Bevir and Blakely, 2016: 38). To construct outcome variables, they are stripped of their contextual and historical specificity.

The TFP evaluation was based on a range of variables to assess whether or not the programme worked (Bewley et al, 2016: 66). One aspect of the evaluation was to consider how inputs from the intervention (the independent variables) were correlated with average changes in the percentage of children from the 'troubled families' who had been permanently excluded from school (an outcome variable). In order to construct these variables there was a process of reification, as the intervention was framed without any contestation about the type of support that was being provided. There was also a process of essentialism; for instance, the child under assessment becomes essentialised as an 'excluded child', without the contextualised knowledge to inform why they were excluded - such as issues of disability, poverty, poor housing or hunger that might have caused the child to struggle in school. Establishment forms of evaluation thus tend to focus on what Chambers (2010: 17) identifies as:

> '...the universal rather than the local; the simplified rather than the complex, the uniform rather than the diverse, the controlled rather than the uncontrollable, the stable rather than the dynamic, the predictable rather than the unpredictable.'

The establishment model of evaluation builds on a 'scientific-logical discourse', which reduces subjective agents into 'pre-determined universal categories' (Krumer Nevo and Sidi, 2012: 302), such as 'troubled families' or 'workless families'. The creation of universal categories through the

construction of variables most often reflects the politics that have framed the policy 'problem' (Taylor, 2006). As such, the construction of criteria for measurement is 'constituted through acts of power' (Connolly, 1995: 128). Participants have little scope to question the broader framing or construction of variables, while evaluators who have been commissioned to provide an assessment of a policy rarely question these pre-determined foundations (Porter, 2015: 248). The original framings of a policy are uncritically followed through the very core building blocks of measurement for the evaluation - thus setting the boundaries of investigation within which the inquiry proceeds.

Privileging pre-determined relations of cause and effect and the 'noise' of context

Evaluation with an establishment orientation follows the original framing of a programme, focusing on the particular relations of cause and effect that have been anticipated by the policy-makers. Social causation is often much more complex than the regularity conception accounts for. There are so many different possible effects happening in the social world around us that it is simply not possible to control for them all (House, 2001). If the causation is complex, then the regularity conception struggles to produce definitive conclusions. Regression analyses are most often used to estimate the average causal effect of one or more outcome variables (Vis, 2012: 169-171). Regression is useful to be able to estimate the net effects of a particular causal pathway (Vis, 2012: 190), but co-variational analysis techniques, such as regression, run into problems when there are too many factors that can influence an outcome and as such find it more difficult to calculate the contribution of each factor.

Complexity is dealt with by aiming to isolate the different causes and effects from other contextual 'noise'. The linear relationship of cause and effect originally anticipated through the policy design is privileged. Evaluators implementing experimental and quasi-experimental research designs often include different methods that aim to deal with complexity. Qualitative approaches can be used to triangulate findings and to explain the differences in outcomes experienced by participants (Bamberger, Rao, and Woolcock, 2010; Haymand, 2013; Pierre, 2004; Reichardt and Mark, 2004; Stern et al., 2012 cited in Moore et al 2018). But what is significant to note is that such data tends to be used through establishment evaluation to explain the limitations and 'contaminations' of their study (Moore et al, 2018: 9-10). Data on complexity is very rarely used to improve the capacity for learning about dynamic and unexpected interactions that happen within the systems in which the intervention is situated.

In the 'troubled families' example, individualised criteria such as 'worklessness', school exclusion and crime or ASB were considered but structural factors such as class, gender, and racial inequalities were not; nor were other issues such as labour market conditions, housing quality and supply, household income or welfare reforms. The first phase outcome of 'moving off out-of-work benefits and into continuous employment' did not take into account the type of work that was secured, or the possible impact that low-paid, poor quality or insecure work may have on family life. Similarly, the desire by the government to see school attendance improve did not necessarily seek to improve the school experience for the child, and there is no evidence of concern for any learning that did or did not take place once attendance had been registered. Such issues were outside of the way the policy had been constructed and were considered to be outside of the boundaries of investigation for the TFP case is not the first time that evaluations revolving around the 'family intervention' model have been accused of examining the approach without considering the wider context. Garrett (2007: 204) highlighted the 'forgetfulness of social policy research' relating to the original 'pioneer' FIPs of New Labour and argued that:

> 'social policy researchers, frequently funded by government grants when investigating "social problems", need to retain a certain wariness and scepticism – a reflexive hesitancy – before providing research "products" which seem to largely endorse policy and practice "solutions" that the State, never independent from class relations, has formulated.'

The demands of an establishment evaluation based on evidencing predictive causal relationships to demonstrate whether an intervention worked as intended or not means that choices need to be made about which factors should be privileged, and which aspects should be screened out. All evaluations are in some way 'bounded' through the selection of certain criteria of performance, and this is very obviously the case for 'troubled families'. These boundaries define the frames of investigation that will take place. Boundaries are established according to what is considered necessary to be able to judge the value of an intervention. Choices must be made to determine the boundaries that frame the investigation of an evaluation. The boundaries for establishment forms of evaluation tend to be pre-determined by commissioners of an evaluation (Picciotto and Donaldson, 2016: 72). Data that is produced outside the boundaries of the primary frames of concern (such as the income level of families involved with the TFP or the effects of welfare reform) are deemed to be less relevant – and so are excluded to enable causal analysis (Ulrich, 2000). Explanations that are consistent with the original framing of policy are privileged over others (Standring, 2017: 232). There is limited scope for contextual data about socioeconomic structures,

symbolic structures or the context of place and time to unsettle the pre-determined selection of performance measures. As a result, the potential for learning beyond original frames of the policy are limited.

In the TFP first phase evaluation, one might hope to have read about issues such as poverty, austerity and welfare reform, as well as their potential consequences in the report on family experiences and outcomes. The TFP evaluation studies showed little evidence of dialogue about the broader structural factors that affected the families. Despite the TFP allegedly working with nearly 120,000 families in the first phase, only 22 families were interviewed in total and only eight were interviewed at both the beginning and the end of their intervention. The second phase evaluation contains no qualitative research with families on the programme. Consistent with establishment forms of evaluation, there were few opportunities for families 'to discuss, explain, justify or defend the meaning, intention or contradictions of their quantified behaviours' (Durose and Richardson, 2016: 26). As a result, the evaluation reports do not tell us whether families believed their lives had been 'turned around' by the programme, whether they thought the programme had helped them make 'significant and sustained progress', how social inequalities limited the possibilities for positive outcomes, or what they thought of being labelled as 'troubled families'.

In the qualitative research with families in the first phase of the programme, background narratives and issues experienced by families prior to intervention are presented as problems of the family; the material circumstances of the families, including the quality of housing and the level of their income, are presented as 'compounding', rather than contributory factors:

> Families often recalled having complex needs that extended back over a period of many years, and sometimes, crossed several generations of the same family. The problems varied according to each family's circumstances but commonly included a combination of physical or mental ill health, drug and alcohol misuse, and educational and behavioural problems affecting their children. *Their situation was often compounded by living in poor quality conditions, and a low income* (Blades et al, 2016: 14 emphasis added).

The interviews with families instead focus on how the intervention has helped them, and how the expert 'family workers' have assisted them - this was, after all, the expected solution to the 'problems' of the 'troubled families' as originally framed through the policy. It is this anticipated

causal relationship between the family intervention approach and families being 'turned around' that the evaluation set out to evidence. The anticipated causal relationship between the input of the intervention and the output of 'turned around' families was privileged above any other, despite 20 per cent of families in the second phase survey not being able to remember the name of their key worker (Ipsos Mori, 2019: 91). The pre-determined evaluation framework privileged investigating whether the intervention delivered according to the original framing of the TFP and the impacts that it was expected to have. Other factors, outside of how the policy was originally framed, were largely excluded. The evaluation process failed to include, or even recognise, the expertise of many of the families in managing very difficult circumstances, often made worse by the services and agencies that were tasked with helping them. In doing so, it mirrored the 'expert' approach of the programme, where technocratic knowledge is privileged. Such a decontextualised approach is consistent with Vanderplaat's (1995: 87) explanation that:

> '...a particular aspect of everyday life is discursively isolated from its experiential context and conceptually reorganised to fit the instrumental capacities of the state administrative apparatus... this discourse is based on the belief that professional science can reduce the muddle of everyday phenomena to empirically identifiable and systematically actionable occurrences.'

The particular contexts of participants, which could add more layers of understanding for the TFP evaluation, are not accounted for as they do not fit within the pre-determined boundaries of concern that are identified through the policy: namely, that intensive and assertive family intervention can change the putative, problematic behaviours of disadvantaged families without recourse to structural modifications. As such, learning from the evaluation is limited to the frames in which the policy was originally conceived. The boundaries of investigation for an establishment evaluation are determined by the original frames of policy. The definition of these boundaries is inherently political (Schwandt, 2015, p.463) and the boundaries of investigation are rarely questioned through establishment forms of evaluation (Picciotto and Donaldson, 2016, p.70). Such an uncritical approach is found because the purpose of establishment evaluation is to produce knowledge that can be used by policy-makers to assess how an intervention has delivered against its original aims. While this is a reasonable approach to take, it limits the scope for learning within existing frames of social policy and, as a result, fails to realise the critical potential of evaluation.

Opening up the boundaries of investigation to critique the original framings of policy

Establishment evaluations rarely question how policy 'problems' have been constructed, operating instead within the problem-solving model of policy-making (Weiss, 1977: 544). Through the problem-solving model, interventions and outcomes are often clearly defined at the start of the process to 'solve' a particular policy 'problem' (Byrne, 2013). Becker (1966: 10) noted that the definition of a problem usually contains, either implicitly or explicitly, suggestions for solutions. The 'problem' is thus constructed as though it has already been 'solved' hypothetically (Turnbull, 2006) and the anticipated causal relationship is then tested to determine whether the intervention 'worked' or not. The problem-solving model 'accepts the world as it finds it, and thus accepts the status quo as the framework for action' (Cox cited in Lingard, 2013: 120).

Turnbull (2006) identifies three main issues with the problem-solving approach: firstly, that complex problems are defined as though they are value-free; secondly, that the model privileges solving these pre-defined 'problems' rather than contesting the way problems are set; and thirdly, that the scientific process is used to 'disguise and/or suppress' political disagreement. Through adopting an uncritical stance towards the construction of a policy 'problem', establishment evaluation contributes to the depoliticisation of policy-making (Turnbull, 2008: 84). Strassheim (2017: 240) argues that the problem-oriented approach produces knowledge that shifts 'politico-epistemic relevances in predetermined directions while creating zones of ignorance'. The evaluations of the TFP reflect the construction of the policy, in which structural explanations of poverty remain relegated to the silences of policy discourse. That the evaluations mirror these omissions and silences highlights the role that 'independent' evaluators can play in producing and reproducing the work of the state and the framing of problems - and problem groups - that it decides upon (Bourdieu, 1996). The boundaries for investigation are pre-determined by the original framing of the policy, thus limiting the scope for producing knowledge that might critically question the status quo.

Opening up the boundaries of investigation to dialogue creates the potential to to problematise how a policy problem has been framed. The practice of boundary critique (Midgely, 1992; Ulrich, 1996; Ulrich and Reynolds, 2010) adopts a critical perspective. Boundary critique identifies how boundaries of investigation are socially constructed and are the result of selective and partial decisions. Boundary critique involves asking questions about the sources of motivation, control,

knowledge and legitimacy that shape the frames of investigation in order to aid reflective practice and question boundary judgements that have been made (Ulrich, 1996: 44).

Through opening up boundaries to potentially dissenting narratives, the process of evaluation can become more critical and dialogic. Ulrich and Reynolds (2010: 243) present boundary critique as a 'participatory process of unfolding and questioning boundary judgements rather than as an expert-driven process of boundary setting'. Unfolding the possibilities of what might be included in an evaluation makes the process more open and accountable to local contexts. Ulrich and Reynolds argue that the 'process of decision-making should make transparent the boundary judgements on which the claims to be decided upon rely' (2010: 263), suggesting that evaluators should highlight alternative boundary judgements and then examine the consequences of the given proposals in light of alternatives and other options. Churchman demonstrates that the boundaries of analysis are crucial and that 'something that appears to be an improvement given a narrowly defined boundary may not be seen as an improvement at all if the boundaries are pushed out' (Midgely et al, 1998: 467).

Schwandt (2018: 134) argues that although 'evaluative thinking is commonly conceived as an individually managed and performed, analytic practice emphasizing sound reasoning and argumentative lucidity about judgments of value', it is also a 'deliberative, practical, reflective, collaborative practice' that can include multiple stakeholders and interested parties in the framing of both problems and solutions. In this process, the questions of what we should investigate, what should be included in the framing of problems, and what should be included or excluded become critical dimensions through which evaluation can challenge and critique - opening up new ways of learning about social policy. Boundary critique is therefore a social practice that 'sweeps in' the views of as wide a range of stakeholders as possible:

> Reflection on boundaries enables consideration of who and what is included, excluded or marginalized and the ethical consequences of such actions. Hence, engaging in boundary critique… offers evaluators, stakeholders, and citizens a means to voice their dissent in ways that re-define and reframe conversations about where we are, where we ought to be going, and what makes that choice the right way to go (Schwandt and Gates, 2016: 70).

Boundary decisions include ethical, moral and normative dimensions and are not merely empirical, methodological or instrumental concerns (Schwandt, 2015). Boundary critique should not ignore the original aims of an intervention, but ensure that boundary setting is transparent, open and

dialogic (Schwandt, 2015: 463). Boundary critique opens up possibilities for alternative narratives to be able to contest the original framings of policy.

Creating opportunities for boundary critique disturbs the privileged position of evaluators (and their commissioners or funders) over other stakeholders in the evaluation process. Through the co-construction of boundaries to frame an investigation for evaluation, the process becomes more of a 'collective undertaking rooted in social interaction and negotiation and situated in [local] contexts' (Schwandt, 2018: 131). Evaluation thus becomes a process of dialogue and critical assessment that includes contextual data in the unfolding of evaluative explanation. Expertise can be found in different places, with different perspectives being offered on the value of decisions and priorities that have shaped an intervention. Through a more contextualised and dialogic approach, we can learn about the particular socioeconomic structures, symbolic structures and the contexts of place and time that affect the lives of people who have been through an intervention, and, in doing so, broaden the scope for learning about how it has worked and what the limitations might have been. Boundary critique developed through dialogue can re-shape the frames of investigation, opening up the potential to engage with alternative explanations about the value and effectiveness of an intervention, situating these within the contexts that give them meaning.

Contextualised evaluation for critical social policy: a discussion

We have argued, using the illustration of the TFP, that establishment evaluation excludes contextual data about socioeconomic structures, symbolic structures and the context of place and time; this means that the pre-determined boundaries of evaluation that are shaped by dominant discourses and frames of social policy are left undisturbed. As a result, the 'role of discovery and invention' is suppressed (Turnbull, 2006:9). If complex social issues are only seen through one particular lens, then the capacity for learning is limited (Chambers, 2015). Decontextualised evaluation set within pre-determined boundaries limits the potential for learning beyond these frames.

A more critical form of social policy requires an alternative approach to evaluation that opens up the boundaries of investigation to include evidence on particular experiences of socioeconomic structures and symbolic structures that can provide us with a greater understanding of participant's lives and the conditions under which a programme is being implemented. Through a more contextualised form of evaluation that can open up new boundaries of investigation, there are many factors that can influence an outcome, and many different possible outcomes. An evaluator should

not neglect these as they might be just as important for explaining the effects of an intervention as the original intentions (Porter, 2015).

An anti-naturalist foundation for evaluation can create more scope to learn beyond the original framing of a policy by opening up space to multiple narratives to create new boundaries of investigation - without abandoning the aim to explain cause and effect. The basis of explaining a causal relationship through an anti-naturalist approach changes from an '"effects-of-causes-stance" (a 'what works' question) which is typical for establishment evaluations, to more explanatory "a causes-of-effects stance" (why does it, or did it not work?' (Pattyn et al, 2017: 7). Factors such as alternative narratives, historical location, or socio-economic contexts are included within analysis due to the belief that actions can only be understood if they are situated in their particular subjective explanations and contexts (Bevir and Kedar, 2008: 507). Methods associated with a naturalist position are not abandoned: rather, they are treated as a source of knowledge to be assessed in dialogue within broader contexts of meaning (Byrne, 2002). Causal explanation with an anti-naturalist foundation is narrative, historical and contingent (Bevir and Blakely, 2016: 34). Causal explanations therefore become situated in the contexts that give them meaning (Bevir & Kedar, 2008, p.508). As different narratives are potentially valid, an evaluation must draw on evidence to construct an explanation and demonstrate why it is preferable to others.

More contextualised causal analysis must be dynamic and informed through dialogue with research participants who have their own accounts of the social world (Bevir and Kedar, 2008: 512). Boundary critique provides the basis for disrupting how a policy intervention has originally been framed and enables reflection on how the criteria for judgements have been constructed (Picciotto and Donaldson, 2016: 72). Evaluators have no 'in-principle advantage' in determining the criteria for judgement (Ulrich, 2000: 8; Schwandt, 2018: 133). As such, the 'identification, clarification, and application of defensible criteria' for evaluation (Fitzpatrick, Sanders and Worthen, 2004: 5) can be developed through dialogue with participants to include the contexts that shape their everyday lives and offer new perspectives on an intervention. Through situating the narratives of participants in the challenges of everyday life, particular life histories, and alternative explanations, a more 'live' form of evaluation can happen. Contextual factors can broaden the boundaries of investigation, and therefore become a critical source of knowledge to better understand the effects of an intervention.

An evaluation becomes more meaningful when it is connected to particular experiences, rather than abstracted variables designed to capture the 'essence' of universal experience. Haraway (1988,

p.586) argues against technocratic objectification that abstracts meaning but aims to do so without succumbing to the relativism that can happen when socially constructed meanings are abstracted away from embodied materiality. Haraway (1988) makes a compelling argument for 'positioned rationality', which is intended as a scientific approach against both subjective relativism and objective universalism. Through positioned rationality, scientific meaning can be explained through 'situated epistemologies'. Haraway makes the case that subjective narratives must be accountable to the objective conditions in which they are situated. Through Haraway's idea of situated knowledges, the narratives of participants become more meaningful when they are made accountable to their objective location in the social world. Critical theories can be used to support this process of accountability.

Opening up the boundaries of judgement does not dismiss the knowledge of policy-makers or evaluators. Rather, it is a demand to abandon the idea of a 'privileged position' in determining the final evaluative judgements about an intervention (Vanderplaat, 1995: 91). Critiquing existing boundaries and co-constructing new boundaries in the TFP could involve asking families views on, amongst other things: the impact of the programme and the extent to which it had 'turned around' their lives; the role of the key worker; the origins of the issues that made them targets of the programme; what their priorities were and what needed to change to improve their lives; and what were the sources of support in addressing their everyday challenges. The views of practitioners and others involved in the delivery of the programme could be sought on issues relating to the privileging of getting adults 'off out-of-work benefits and into continuous employment' or the exclusion of criteria relating to housing issues and household income when assessing the effectiveness of the intervention. Other stakeholders in the boundary critique process would potentially involve social policy academics and other professionals working with marginalised families.

Boundary critique opens up a more sociological and dialogic approach to policy evaluation. A more contextualised form of evaluation can broaden the scope of learning beyond the original framing of a policy intervention. Such an approach with 'troubled families' would recognise the knowledge, expertise and capabilities of many families in dealing with the vicissitudes of everyday life, including those caused by the government claiming to be helping them via the TFP. Through boundary critique, more contextualised evaluation can be used to open up the frames of investigation for an evaluation, without abandoning the aim to evidence causation. More contextualised evaluation can continue to fulfil its purpose of assessing the impact of a policy intervention and identifying the

cause of these changes but, in addition, it can contribute towards a more critical questioning of how policy 'problems' have been constructed, and provide an architecture to learn about uncertainty and injustice.

References

Becker H (1966) *Social Problems: A Modern Approach*, New York: John Wiley and Sons.

Bevir M. and Blakely J (2016) Naturalism and Anti-Naturalism. In: Bevir M and Rhodes RAW (eds) *Routledge handbook of interpretive political science,* New York: Routledge, pp.31-44.

Bevir M and Blakely J (2017) Why political science is an ethical issue. *Political Studies* 66(2): 425-441.

Bevir M and Kedar A (2008) Concept formation in political Science: an anti-naturalist critique of qualitative methodology. *Perspectives on Politics* 6(3): 503-51.

Bewley H, George A, Rienzo C and Portes J (2016*) National Evaluation of the Troubled Families Programme: National Impact Study Report Findings from the Analysis of National Administrative Data and local data on programme participation*, London: DCLG.

Blades R, Day L and Erskine C (2016) *National Evaluation of the Troubled Families Programme: Families' experiences and outcomes*, London: DCLG.

Bourdieu P (1996) On the family as a realized category, *Theory, Culture and Society*, 13(1): 19-26.

Byrne D (2002) *Complexity theory and the social sciences: An introduction*. New York: Routledge.

Byrne D (2013) Evaluating complex social interventions in a complex world. *Evaluation* 19(3): 217-228.

Casey L (2013) Former Portsmouth police officer back on ASBO beat, *BBC News*, 2 September. Available at www.bbc.co.uk/news/uk-england-hampshire-23896776 (accessed 11 March 2017).

Chambers R (2010) Paradigms, poverty and adaptive pluralism. *IDS Working Paper 344.*

Chambers R (2015) Inclusive rigour for complexity. *Journal of Development Effectiveness* 7(3): 327-335.

Connolly W (1995) Twilight of the idols. *Philosophy and Social Criticism* 21(3): 127-137.

Cook C (2016) Troubled Families report 'suppressed', *BBC News online*, 8 August. Available at http://www.bbc.co.uk/news/uk-politics-37010486 (accessed 17 March 2017).

Crossley S (2018) *Troublemakers: the construction of 'troubled families' as a social problem*, Bristol: Policy Press.

Day L, Bryson C, White C, Purdon S, Bewley H, Kirchner Sala L and Portes J (2016) *National Evaluation of the Troubled Families Programme: Final Synthesis Report*, London: DCLG.

DCLG (2013) *Massive expansion of Troubled Families programme announced,* 24 June. https://www.gov.uk/government/news/massive-expansion-of-troubled-families-programme-announced (accessed 12 March 2017).

DCLG (2017) *Supporting disadvantaged families Troubled Families Programme 2015 to 2020: progress so far*, London: DCLG.

Durose C and Richardson L (2015). *Designing public policy for co-production*. Bristol: Policy Press.

DWP (2017a) *Improving Lives Helping Workless Families*, London: DWP.

Ecorys (no date) *Policy and Research*. Available at https://uk.ecorys.com/viewservice/14 (accessed 25 April 2019).

Everitt A and Hardiker P (1996) *Evaluating for good practice*. Basingstoke: Macmillan.

Fitzpatrick J, Sanders JR and Worthen BR (2004) *Program Evaluation: Alternative Approaches and Practical Guidelines* (3rd ed.) Boston: Allyn and Bacon.

Freeman M and Vasconcelos EFS (2010) Critical social theory: Core tenets, inherent issues. In M. Freeman (Ed.), Critical social theory and evaluation practice. *New Directions for Evaluation* 127: 7–19.

Gardiner M (2010) Everyday utopianism: Lefebvre and his critics. *Cultural Studies* 18(2): 228–254.

Garrett PM (2007) '"Sinbin" solutions: the "pioneer" projects for "problem families" and the forgetfulness of social policy research', *Critical Social Policy*, 27(2): 203-230.

Gillies V and Edwards R (2012) Farewell to family? A reply, *Families, Relationships and Societies,* 1(3): 431–4.

Greene J (2009) Evidence as "proof" and evidence as "inkling". In: Donaldson SI, Christie CA, Mark MM (Eds.), What counts as credible evidence in applied research and evaluation practice?Thousand Oaks, CA: Sage, pp.53–167

Greene J (2013) Evaluation, Democracy, and Social Change In I.Shaw, J.Greene and M.Mark. (Eds.), Handbook of evaluation London: Sage, pp.119-140

Guba EG and Lincoln YS (1989) *Fourth generation evaluation*. London: Sage.

Haraway D (1988) Situated knowledges: the science question in feminism and the privilege of partial perspective. *Feminist Studies* 14(3): 575-599.

Hargreaves C, Hodgson P, Mohamed JN and Nunn A (2019) Contingent coping? Renegotiating 'fast' disciplinary social policy at street level: Implementing the UK Troubled Families Programme. *Critical Social Policy*, *39*(2): 289–308.

House E (2001). Unfinished business: causes and values. *American Journal of Evaluation* 22(3):309-315.

Ipsos Mori (2019) Troubled Families Programme National Evaluation Family survey – Follow-up survey, London: Ipsos Mori. Available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/786890/National_evaluation_of_the_Troubled_Families_Programme_2015_to_2020_Follow_up_family_survey.pdf (accessed 25 April 2019).

Krumer-Nevo M and Sidi M (2012) Writing against othering. *Qualitative Inquiry 18(*4): 299-309.

Lingard B (2013) The impact of research on education policy in an era of evidence-based policy. *Critical Studies in Education* 54(2): 113-131.

Martens D and Wilson AT (2012) *Program evaluation theory and practice: A comprehensive guide*. New York: Guilford Press.

MHCLG (2019a) More families in need making significant progress, London: MHCLG. Available at https://www.gov.uk/government/news/more-families-in-need-making-significant-progress?utm_source=0ccd9645-c674-4e66-a6d6-ff1f88557701andutm_medium=emailandutm_campaign=govuk-notificationsandutm_content=immediate (accessed 25 April 2019).

MHCLG (2019b) *Building Resilient Families: Third annual report of the Troubled Families Programme 2018-19*, London: MHCLG. Available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/790402/Troubled_Families_Programme_annual_report_2018-19.pdf (accessed 25 April 2019).

Midgley G (1992) The sacred and profane in critical systems thinking. *Systems Practice* 5: 5-16.

Midgely G, Munlo I and Brown M (1998) The Theory and Practice of Boundary Critique: Developing Housing Services for Older People. The Journal of the Operational Research Society, 49(5): 467-478.

Moore M, Parsons B and Jessup P (2018). How evaluators can use a complex systems lens to get "untrapped" from limiting beliefs and assumptions. *American Journal of Evaluation* 39(1): 1-17.

Pattyn V, Molenveld A and Befani B (2017) Qualitative Comparative Analysis as an evaluation tool: Lessons from an application in development cooperation. *American Journal of Evaluation* 40 (1): 55-74.

Picciotto R (2015) Democratic evaluation for the 21st century. *Evaluation* 21(2): 150-166.

Picciotto R and Donaldson S (2016) *Evaluation for an Equitable Society*. Charlotte: IAP

Porter S (2015) The uncritical realism of realist evaluation. *Evaluation* 21(1): 65-82.

Purdon S and Bryson, C. (2016) *Evaluation of the Troubled Families Programme Technical report: impact evaluation using survey data*, London: DCLG.

Rossi PH, Lipsey MW and Freeman HE (1999). *Evaluation: A systemic approach* (6th ed.). London: Sage.

Sanderson I (2002) Evaluation, policy-learning and evidence-based policy making. *Public Administration* 80(1): 1-22.

Schwandt TA (2007) Expanding the conversation on evaluation ethics. *Evaluation and Program Planning* 30(4): 400-403.

Schwandt TA (2015) Reconstructing professional ethics and responsibility: Implications of critical systems thinking. *Evaluation* 21(4): 462-466.

Schwandt TA (2018) Evaluative Thinking as a Collaborative Social Practice: The Case of Boundary Judgment Making. *Evaluative Thinking*, 158: 125-137.

Schwandt TA and Gates E What Can Evaluation Do? An Agenda for Evaluation in Service of an Equitable Society. In: Donaldson S and Picciotto R (eds) *Evaluation for an Equitable Society*. Charlotte, NC: Information Age Publishing, pp.67-82.

Strassheim H (2017) Bringing the political back in: reconstructing the debate over evidence-based policy. A response to Newman. *Critical Policy Studies* 11(2): 235-245.

Stake RE (2000) A modest commitment to the promotion of democracy. *New Directions for Evaluation*, 85: 97-106.

Standring A (2017) Evidence-based policymaking and the politics of neoliberal reason: a response to Newman. *Critical Policy Studies* 11(2): 227-234.

Taylor D and Balloch S (2005) *The politics of evaluation: Participation and policy implementation*. Bristol: Policy Press.

Taylor D (2006) Critical policy evaluation and the question of values: a psychosocial approach. *Critical Social Policy* 26(1): 243-267.

Turnbull N (2006) How should we theorise public policy? Problem solving and problematicity. *Policy and Society* 25(2): 3-22.

Turnbull N (2008) Harold Lasswell's "problem orientation" for the policy sciences. *Critical Policy Studies 2*(1): 72-91.

Ulrich W (2000) Reflective practice in the civil society: the contribution of critically systemic thinking. *Reflective Practice* 1(2): 247-268.

Ulrich W (1996) *A Primer to Critical Systems Heuristics for Action Researchers.* Hull, UK: University of Hull, Centre for Systems Studies.

Ulrich W and Reynolds M (2010) Critical Systems Heuristics. In M Reynolds and Holwell S (Eds.) *Systems approaches to managing change: a practical guide*. London: Springer, pp.243-292.

Vanderplaat M (1995) Beyond Technique. *Evaluation* 1(1): 81-96.

Vis B (2012) The comparative advantages of fsQCA and regression analysis for moderately large-N analyses. *Sociological Methods and Research* 41(1): 168-198.

Weiss C (1977) Research for policy's sake: the enlightenment function of social research. *Policy Analysis* 3(4): 531-545

Weiss C (1979) The many meanings of research utilization. *Public Administration Review* 39(5): 426-431.

Weiss C (1993) Politics and evaluation: a reprise with mellower overtones. *Evaluation Practice* 14(1): 93-106.