eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

The Troubled Families Programme: Learning about policy impact through realist case study research.

Sadie Parr, Research Fellow, CRESC, Sheffield Hallam University S.Parr@shu.ac.uk

Harriet Churchill, Lecturer in Social Work, Department of Sociological Studies, University of Sheffield h.churchill@sheffield.ac.uk

Abstract

This article aims to critically explore how qualitative case study research that is founded on realist principles can fundamentally enhance social policy evaluation methodologies and, in turn, provide improved learning for policy makers and practitioners. We suggest these methodological advantages are accrued through the careful construction of theory-based explanations of 'how' policy programmes work thereby addressing the limitations of quasi-experimental methods - namely a focus on and prioritisation of outcome measures. The paper situates this key argument within wider, long-standing debates about evidence-based policy-making and what constitutes 'evidence' of impact in social policy. It does so through reflection on the contentious and contradictory knowledge claims that surround the Troubled Families Programme and evaluative claims regarding its efficacy. In conclusion and looking forward, we suggest that there remains much scope to combine 'intensive' qualitative case studies with 'extensive' quantitative measures within local and national evaluations of complex, multi-dimensional social policies, such as the Troubled Families Programme.

Key Words Troubled Families; critical realism; evaluation; case study

Introduction

There has been nearly 20 years of research in the UK, including several official evaluations, concerned with evaluating what are commonly referred to as 'family intervention services' (FIS). Yet there is little consensus on what the evidence base tells us, with distinctly opposing views among and between policy makers, practitioners and academics. Most recently, highly charged debates concerning the high profile Troubled Families Programme (TFP) have added fuel to these opposing positions. While there have been important and robust critiques of the TFP on ideological grounds (Crossley, 2018), much of the controversy surrounding it has oscillated around the nature of the 'evidence' that informed the policy programme and assessments made regarding its efficacy. The much publicised research findings that emerged from the independent evaluation of phase one (spring 2012 to spring 2015) of the TFP (Day *et al*, 2016), and found it wanting, has added another significant dimension to this evidence debate.

The TFP provides a particularly compelling example through which to bring into focus longstanding debates about the nature of different kinds of knowledge claims and what constitutes 'evidence' of impact in social policy. In this paper, we draw attention to the continued predominance of quasi-experimental methods within evaluation methodologies and the shortcomings of this approach. We are centrally concerned with the ways in which methodologies, often based on the prioritisation of quantifiable outcome measures, have limitations in terms of explanatory value and middle-range theorising, which inhibits potential policy and practice development. We reframe the methodological debate about 'evidence' by examining the role of case study research in social policy evaluation. Despite this approach being commonplace in evaluation methodologies, we suggest that the explanatory potential of case studies has not been fully realised. We critically explore the

key claim that case study methods have the potential to open up the 'black box of implementation' and generate context-sensitive theorising that extrapolates 'mechanisms of change' in order to better understand outcomes.

The first section of the paper presents an overview of the policy context and the gathering momentum in favour of the adoption of FIS models which informed the TFP. The second section summarises the key findings from formal evaluations of prior FIS initiatives which formed a core part of the 'evidence' base for the TFP. This section also reviews the critical commentary that surrounded these knowledge claims (Crossley, 2018; Gregg, 2010; Garrett, 2007; Fletcher *et al*, 2012). We then discuss the evaluation of the first phase of the TFP (2012-2015) in section three, raising critical issues about its methodological approach. In section four, we situate these reflections within wider debates about evidence-based policy making (EBPM) before arguing in section five for greater emphasis on multi-method theoretically generative evaluation research that includes realist case studies. We suggest there remains much scope to combine 'intensive' qualitative case studies with 'extensive' quantitative measures within local and national evaluations of complex, multi-dimensional social policies such as the TFP.

The Troubled Families Programme: the social, political and historical context

The TFP was launched by the UK Coalition government in late 2011 following the riots that took place in cities across England months earlier. Through a focus on the failings of families and parents, a narrative of blame and individual deficit and culpability attributed the problems of rioting, minimally to structural issues of poverty and inequality, and mainly to behavioural causes (Cameron, 2011). This narrative gave impetus to the TFP in which the 152 upper tier local authorities (LAs) in England were expected to 'turn around' the lives of

3

an estimated 120,000 families in order to help them out 'of this cycle of despair and give their children a better chance in the future' (DCLG, 2014: 6).

£448 million of cross-department central government funding was allocated to the programme with the Department of Communities and Local Government (DCLG) having the lead responsibility for the programme. The DCLG allocated much of this funding to LAs tasked with 'turning around' a specified number of 'troubled families' by 2015. While LAs were provided with capital up-front to help establish the programme, funding was mainly provided on a payment by results (PbR) basis whereby LAs received a payment for engaging each family to the programme and further funding for families achieving specified 'outcomes'. Rather than prescribing a definitive model of intervention, the TFP largely constituted a national framework of data/outcomes governance and funding targeted at families with specific co-occurring 'problems' namely: household welfare reliance; school exclusion, truancy and persistence school absence problems; and youth convictions or youth and/or adult anti-social behaviour problems. Beyond this, LAs were to decide service provision priorities and delivery approaches. Central government guidance, however, advocated "the family intervention approach" as "the most compelling method of intervention" (DCLG, 2012a: 4).

Controversially, by June 2013 the Coalition announced the TFP would be significantly expanded. The 'Extended TFP' has been in operation since early 2015 and funded until spring 2020. It targets a larger and more diverse group of 'troubled families' (an additional 400,000) including those affected by severe health problems, domestic abuse, family conflict, household debt and 'children in need' as defined by the 1989 Children Act. The expanded programme was also accompanied by a revised financial framework while the terminology of

'significant and sustained progress or continuous employment' replaced the 'turned around' terminology of phase one (Bates and Bellis, 2018).

Although the national framework for the TFP is distinctive, the 'family intervention approach' reflected strong similarities with New Labour Government's 'family intervention projects' (FIPs) introduced in the early 2000s (Flint *et al*, 2011). This national network of FIPs had been preceded and inspired by a number of longer-established local initiatives including the well-known Dundee Families Project (DFP). What all of these projects ostensibly had in common was a commitment to a model of working premised on a key worker providing intensive support to a family over a prolonged period, taking a 'whole family' approach, appraising need, providing one-to-one support but also procuring and co-ordinating the work of other services and agencies.

The scale and speed with which family intervention as a broad model of working has been adopted across England since the late 1990s, along with the substantial resources committed to it, is remarkable. The reasons for the adoption of the approach are multifaceted, including the prevailing political climate, ideological commitment, financial imperatives, public opinion as well as 'evidence' (Cairney, 2019). We are concerned here with the latter and with the status of evidence derived from evaluation research specifically.

Claims and counter claims

A key part of the evidence claims that informed and justified the TFP derived from independent, formal evaluations of the family intervention pilots and programmes introduced under New Labour (see Dillane *et al*., 2001; Nixon *et al*., 2006 and 2008; Jones *et al*., 2006). These evaluations included both qualitative and quantitative methods and measures of success as well as cost-benefit analyses. While mindful of limitations, such as in relation to

5

longer-term outcomes and the extent to which more traditional family support might achieve similar results, the authors of these evaluations pointed to significant positive outcomes for between 59% (Dillane *et al*, 2001) and 80% (Nixon *et al*, 2006) of families. They also reported affirmative views of the support from a range of 'stakeholders', together with cost savings for the tax payer: "There is general consensus that the FIP model is 'fit for purpose' and is required to deal with the families they are targeting" (White *et al*, 2008: 146). Despite important differences between the projects, the evaluations likewise identified several common factors and practice principles associated with positive outcomes (Flint *et al*, 2011).

For the TFP, the DCLG (2012a) claimed to have brought together this academic evaluative evidence yet arguably promoted a 'diluted, limited' view of services and practice (Crossley, 2018). While some limitations to the evidence-base were recognised – such as the absence of studies that have employed control (or comparison) groups and a lack of studies incorporating outcome assessments based on 'objective external data sources' – it was concluded "there is compelling evidence that it works" (DCLG 2012a: 11). The DCLG side-lined some evidence however and overlooked important caveats to evaluation findings as well as marginalising more critical 'conceptions of the problem' (Crossley, 2018; Ball *et al*, 2012). Many families' difficulties were disregarded while others overstated; poverty was never mentioned (Sayer, 2017; Hayden and Jenkins, 2014). In contrast to our discussion of case study research below, official TFP publications also tended to include highly anecdotal summarises of 'family cases' to support the stark deficit view of 'troubled families' and the ways in which family intervention workers can facilitate positive changes (e.g. DCLG 2012b).

While framed as 'strong' by policy makers, the evidence base to support the broader parameters and ways of working promoted by the TFP has been widely criticised by others and accused of failing to provide reliable 'evidence'. Crossley (2018) and Garrett (2007), for instance, have contested the positive FIP evaluation claims drawing attention to the ambiguous nature of 'success' in the DFP evaluation, highlighting the number of families re-referred in the sample, the potential for accepted referrals to have been cherry picked and the dissatisfaction of some families and stakeholders. The evaluation by Nixon *et al* (2006) was critiqued on grounds of sample size, misleading findings that measured impact primarily for those families which 'engaged', and an apparent over-interpretation of data which purportedly overlooked important facets of families' narratives. According to Garrett (2007: 223) the evaluation was 'far too buoyant and emphatic' which in turn legitimated policy and programme inadequacies.

Writing from a different standpoint, one that privileges rigorous quantitative measures and evidence of impacts, Gregg (2010) strongly critiqued the claims of 'success' reported in some of the FIP evaluations. For him, the evaluations were flawed because measurement of 'behaviour outcomes' were limited, findings were 'opinion-based' (qualitative), with small sample sizes (biased towards those who engaged) and no control groups. Fletcher *et al* (2012) also reported grave concerns about the limited scope, scale and quality of UK-funded and conducted programme evaluation research and systematic reviews. In contrast to endorsing FIS as 'evidence-based' and therefore by implication desirable and cost-effective, they raised concerns about the potential for harmful impacts, in part due to their limited evidence-base and therefore by implication potential poor and varied quality. They also point to the potential risks associated with key features of their design, in particular, their use of 'assertive, coercive and conditional' modes of engagement and delivery which endorse statutory and professional dominance over families and potentially restrict family and social

7

rights. The TFP has been consequently depicted by critical commentators as an archetypal example of 'policy based evidence' whereby policy makers' prior agendas construct the problem and shape the preferred solution and 'evidence' operates to support these decisions (Cairney, 2019; Gregg, 2010).

The TFP evaluation: a narrative of failure

The unease among critical commentators regarding the evidence on which the TFP has been legitimated has intensified as the programme has evolved. An early source of dispute was its initial claims of unprecedented success. Alongside the crude use of 'family case' stories, 'policy success' for the first phase of the TFP was achieved when family-level data collected by LAs indicated specified problems within families had reduced to specified levels over specified and short-term timeframes. These 'outcomes' were taken to provide evidence that the lives and prospects of troubled families had been 'turned around' (NAO, 2016). In June 2015, according to DCLG data (NAO, 2016), the programme had succeeded in 'turning around' 99% of the families worked with, which Crossley (2015) critiqued as 'unbelievable' success – the TFP appeared to be 'a perfect social policy'. Not only does 'turned around' overstate the kinds of changes families might have made and fails to recognise the fluid, fragile and non-linear nature of change within families but the financial incentives under the PbRs system led LAs to demonstrate family outcomes via data matching exercises rather than the delivery of any actual TFP intervention (Crossley, 2018).

The national evaluation of phase one of the TFP (Day *et al*, 2016) is the most recent key development in the debate about the efficacy of the policy[1] which, in comparison to the above, sought to more extensively evaluate policy and programme impact. As an independent national evaluation study it was based on a more rigorous approach to 'evidence'

---

[1] The evaluation of phase two of the TFP (MHCLG, 2018) is ongoing.

and 'evaluation'. It included: a 'process evaluation' comprising qualitative research with case study LAs and TFP co-ordinators; an 'impact evaluation' which employed a quasi-experimental research design and used outcome data from linked national administrative datasets as well as a large-scale survey of families; and an 'economic' or 'cost-benefit analysis' (CBA) which was designed to draw on data from the other two strands, although the results from the impact evaluation in fact disallowed a full CBA (Bewley et al, 2016; Blades et al, 2016; Purdon et al, 2016; White and Day, 2016).

The study was late in providing evidence, with allegations that publication had been held back for a year, due to the null effect in 'turning around' the lives of families:

> Across a wide range of outcomes, covering the key 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 (Day *et al*, 2016: 69).

Although the evaluation was complex both the press release (NIESR, 2016) and the media reporting that it triggered focussed on specific statements within the report - the evaluation was unable to provide any evidence that any significant impacts for families were attributable to the TFP. In a mainly negative debate in which the TFP was described as a 'monumental failure' (Bonell, 2016) and the evaluation as 'devastating' (Butler, 2016), the positive aspects of the evaluation findings, caveats to the methodology and findings about practise were lost. This is despite the national evaluation reporting more mixed and nuanced findings:

The impact evaluation identified statistically significant impacts on families'
satisfaction with the service; their confidence, and optimism about being able to
cope in the future, compared with a matched comparison group of families (Day *et
al*, 2016: 68).

Although the TFP evaluation presented a more mixed picture of TFP effects than media and
commentator reporting suggested, it fell short of providing an account of *why* the TFP did not
appear to work. This is because the study was predominantly an outcomes-focused
evaluation in the sense that it was centrally concerned with assessing the extent to which
there had been evidence of progress in relation to the objectives that the program was meant
to achieve.   Although the evaluation told us some things about 'process' in addition to the
direction of effects, issues around implementation and context were considered discretely and
not theorised in relation to outcomes (Day *et al,* 2016). One challenge was the TFP
encompassed diverse local approaches. The lack of comparability of these services critically
limited the capacity of the national evaluation to arrive at meaningful outcomes-orientated
assessments at the national level. The evaluation acknowledged it potentially amalgamated
TFP services that may have produced positive outcomes with those that did not:

Averaging effects between local areas may therefore have influenced the impact
results at a programme level - that is, in principle it is possible that positive impacts in
some areas were obscured by negative impacts elsewhere, so average impacts were
not significantly different from zero (Day *et al*, 2016: 77).

Within the remit of the TFP evaluation little analytical attention could be paid to the
circumstances within which positive impacts might have been produced nor the possibility

that family support provision in the comparison groups may have fostered ways of working and outcomes similar to the TFP (Canavan *et al*, 2016). Although the evaluation provided some understanding of the different TFP operating models and how certain practice principles are embedded in practice, the evaluation was not able to impart knowledge about how the specifics of TFP services impact on a diverse range of outcomes. It is also important to note that the outcome-orientated approach measures success in very particular, standardised ways and, in the case of the TFP, largely, according to the dictates of the PbR framework. Such measures are not neutral however but imbued with values and ethical considerations, and only those who regard the intended goals as just and desirable are likely to perceive their achievement as indicative of 'success' or 'effectiveness' (McConnel, 2010).

These caveats have not been fully acknowledged within national policy debates and to some extent the critical literature. It could be suggested that the findings from the TFP evaluation have been used to bolster socio-political critique of intensive family interventions, such that ideological and empirical evaluation have become conflated (Cairney, 2019). There is a certain irony here however that although academic commentators have been keen to dispel 'myths' around the evidence base that informed the TFP, there has been less enthusiasm to mobilise the same methodological and epistemological circumspection to the TFP evaluation and its data collection methods, in order, for example, to offer a more nuanced understanding of the findings. This raises questions regarding what the evaluation tells us about the TFP and what use this knowledge is to researchers, practitioners and policy makers for improving the lives of multiply disadvantaged families. For many, the evaluation has been read as telling us that the TFP represents a policy failure, despite the very different contexts of practice across a complex and diverse social programme.

An uncritical acceptance of the negative programme impact findings within the national evaluation of the TFP also risks reproducing an 'evidence hierarchy' in which quantitative and experimental methods are privileged (Canavan *et al,* 2016). This has ethical as well as methodological implications. Despite a justified concern among critics of the way in which the TFP problematically frames families, evaluations that give value to "subjective assessment" (Gregg, 2010) through listening to the voices of families and other 'stakeholders', are construed as less significant; children and parents are not considered legitimate 'subjects of knowledge' (Skeggs, 1997). This seems somewhat paradoxical. As many commentators have argued, families targeted by FIS and deemed 'troublesome' have a stigmatised and excluded status. In evaluating the success of TFP policy, this 'symbolic violence' (Sayer, 2017) and exclusion is reinforced in the sense that the status of their views are also framed as inferior and their evidence remains largely unheard (Squires and Measor, 2005). In the TFP process evaluation, there was only limited 'snap shot' qualitative research with families positioning their knowledge less worthwhile than standardised and service-based measures.

In several ways in-depth qualitative evidence and service user perspectives have been marginalised in these debates. However, several smaller scale studies have generated and examined the positive sentiments expressed by service users engaged with FIS, including those funded by the TFP (Hoggart, *et al*, 2014; Davies, 2012; Hargreaves *et al*, 2013; Hayden and Jenkins, 2014; Hoggett and Frost 2017). Many of these offer a more nuanced understanding of the instantiation of FIS at a local level. They often also retain a more reflective stance in regard to the TFP and, in particular, the way in which the policy problem, aims, outcomes for families and concepts of 'success' are conceptualised. These have often told positive stories about the impact of intensive interventions. They point to key practice

principles, local agency and evidence of discretion displayed by practitioners, and frequently report positive assessments by families who appreciate at least some of the support provided. Notwithstanding this growing evidence base of small-scale studies of local practice, what collectively these studies might tell us about TFP remains under-examined. It is one example of Ball *et al*'s (2012: 272) concern that the TFP reflects the 'latest instalment in the historical failure to adequately utilise acquired learning'.

What counts as policy 'evidence'?

Questions remain about how to reconcile the contradictory messages arising from large-scale FIP evaluations, the TFP evaluation, and small-scale local service case studies. We address this question through a discussion of the role of the case study in evaluation research. While it is essential to avoid false claims of effectiveness as critics of the FIP evidence base suggest, it is equally important to avoid false or exaggerated claims of ineffectiveness (McConnel, 2010; Tunstill *et al*, 2005). As noted by the national evaluation team, their study "presented a mixed picture with regard to the effectiveness and impact of the Troubled Families Programme" (Day *et al*, 2016: 86). Rather than abandoning policy outright, nuanced understanding of how and why FIS achieve or fail to achieve desired outcomes is required (Houston, 2010). In this section, we consider the ways in which different claims to knowledge can be utilised to expand the evidence base and better achieve cumulative learning. We focus here on the value of case studies for understanding the 'success' of the TFP (and FIS more generally) and in providing useful knowledge for policy makers.

Debates about the nature of the evidence regarding the TFP are of course presaged by longstanding debates about EBPM. The latter is a contested term and much has been written about the nature of 'evidence' and the movement for EBPM, and we do not attempt to summarise in detail the arguments here (see instead e.g. Boaz *et al*, 2008; Weiss, 1997;

Sanderson, 2002). Particularly relevant for this discussion are political and epistemological tendencies to prioritise positivistic approaches to evidence within EBPM. Compared to the former New Labour era (Boaz *et al*, 2008; Sanderson, 2002; Wells, 2007; 2018), in some regards, EBPM has shifted direction with the Coalition and Conservative Governments. However, for several official evaluations the emphasis has remained on the use of quantitative outcome measures and to some extent the valorisation of outcomes evaluation evidence gathered using the 'gold standard' of randomized control trials (RCTs) (Axford and Morpeth, 2013; Haynes et al, 2012). This has been accompanied by a corresponding scepticism towards qualitative research (Wells, 2018).

The TFP evaluation (Day *et al*, 2016) reflected this ambition to develop a 'scientific approach' to evaluating FIS in its attempts to discern 'truths' that would be applicable across the different LAs within which the TFP is operating. The introduction of PbR in public service reform also reflects this focus on quantifiably measured 'success' in the sense that it serves as a real-time monitoring process of a programme's performance (Durose *et al*, 2017). There are methodological counter-arguments, however, regarding the 'right' evidence that should be used to inform policy such that standardised approaches to EBP and evaluative research have been subject to trenchant criticism.

Within the social sciences, since the 1980s, there have been moves to address the short-comings of evaluation methodologies founded on RCT or quasi-experimental designs to improve the range and quality of evidence about complex and dynamic social programmes (Sanderson, 2002., 2009; Marchal *et al*, 2012; Pawson and Tilley, 1997). When social policy programs that include multiple interventions are evaluated, RCT (even if practically possible) or quasi-experimental research studies are not designed to generate extensive evidence to explain how and why outcomes vary according to different contexts of implementation which, 'leaves us none the wiser about where to target resources or how to maximise impact'

(Greenhalgh *et al*, 2015: 1). As already noted, this was the case with the TFP evaluation which failed to provide sufficient insight into *how* the TFP works as a complex, multi-dimensional programme; the 'black box'. The focus was not so much imparting knowledge about how the different initiatives that comprise the programme worked, but providing outcome data on how they performed in relation to pre-defined targets.

These critiques have given rise to an ever-growing debate about what constitutes 'evidence' for the purpose of policy evaluation and development. With that, there has been an increasing drive for what has been broadly called 'theory driven evaluation' which pays greater attention to the causal mechanisms and the contextual factors that generate change (Marchal *et al*, 2012). It is in this vein that case study research has been championed as a way to provide robust contextualised explanation. To date however, while 'theory driven evaluation' (often) in the guise of 'logic models' (or operational programme theory) has gained widespread currency within the evaluation community (Sanderson, 2009; Chen, 1990; Weiss, 1997), the case study, as a broad methodological approach, has failed to gain credibility as a means of providing useful knowledge for policy makers.

It is not the intention to rehearse arguments about the strengths and limitations of case study research (there is a large body of work that offers guidance on case study methodology e.g. Yin, 2014). Suffice to say, that a core argument is that case studies provide in-depth and rich understanding of that which is under study. In social policy evaluation research, case study methodologies focus analytical attention on the diversity and complexity of implementation. Studies of local practice are sensitive to the dynamics that give rise to particular outcomes, paying close attention to the power and role of agency (Welch *et al*, 2011). Given this focus on human agency, case study research will usually take account of the views of those involved in the programme through detailed qualitative enquiry.

The argument that the unique contribution of cases studies is in the provision of rich, and context-sensitive data often goes hand in hand with assumptions that research findings cannot be generalised, and, by implication, have limited value in evaluative studies. While quantitative research is steeped within a positivist tradition in which the external validity of quasi-experimental research designs can be assessed against fairly standard criteria, the validity of case study research has no comparable (universally accepted) quality criteria. It is of course not the case that that qualitative researchers feel that the knowledge they produce is so idiosyncratic that it has no relevance beyond the specific. Certainly many researchers, and particularly those working within policy orientated research and evaluation, would claim that their findings have relevance beyond the particular and can assist in the process of accumulating knowledge of relevance for policy makers. For example, TFP case studies have reported notable similarities in terms of the types of families referred, their circumstances and support needs, the forms of intervention delivered, and outcomes achieved. That said, it is also commonplace for qualitative researchers in their reflective discussions of methodology to promote caution in regards of generalisations (Welch *et al*, 2011).

This reluctance to make a claim for 'generalisation' is because the term invokes the positivist tradition of statistical significance and hypothesis-testing, something at odds with the small sample sizes and the context-specific nature of case study findings. This sentiment is reflected in social policy evaluations in which case studies are fairly common but tend to play second fiddle to quantitative measures reflected in a distinction between 'impact' and 'process' evaluation. In the latter, case study findings often provide limited understanding however by illustrating what has been termed the 'grey box', whereby descriptors of practices might be discerned but their inner workings or ways of operating, and, with that, how they generate change, not fully revealed (Axford and Morpeth, 2013). We suggest that this was the case with the TFP evaluation. At other times, case studies are limited to the initial,

'exploratory' phase of evaluation after which large-scale quantitative testing is undertaken to look for generalisable patterns (Welch *et al*, 2011). Furthermore, when qualitative data from across numerous case studies become subject to (cross-case study) thematic analysis, the context specific nature of findings can also be lost.[2]

While case studies cannot claim to be representative, they can provide more than idiosyncratic understanding. If case studies are to help overcome the limitations of quasi-experimental evaluation designs and help build knowledge about FIS, it is important to be explicit about their explanatory potential, what value they add and how this is to be achieved. This means thinking carefully about how small, localised studies of practice can be used to explain how social policies work 'on the ground'; generate and evaluate 'causal' explanations across contexts and target groups; and consider how explanations can be transferred to different contexts (Ragin, 2014; Welch *et al*, 2011). While this means questioning widely accepted evidence hierarchies it does not amount to endorsing an evidence relativism where one case study is as explanatory as the next (Clegg, 2005). Case study research, while admittedly less precise and formulaic, is by no means a less scientific and robust process.

Theory-driven and mixed method evaluation

In order to better explicate how studies of local practice, fundamentally of context, can be useful for impact evaluation, researchers have drawn on the philosophical insights of realism. Realist inspired case study research involves paying attention to the influence of context by identifying underlying mechanisms of change and how they operate within spatial-temporal contexts (Astbury and Leeuw, 2010). This sort of methodological approach in case study

---

[2] Analysis of local provision is included in the ongoing evaluation of phase two of the TFP but reflecting the point above, the case studies here form part of a 'process evaluation' with the intention 'to better understand the delivery of the programme and to provide descriptive accounts of how the programme is being received by families and delivered by staff' (MHCLG, 2018)

research is explicitly concerned with *how* and *why*, with generating causal explanation. In the parlance of realistic evaluation (Pawson and Tilley, 1997), the aim is to explain 'what works, how, in which conditions and for whom', and so goes beyond and spans the distinction between 'impact' and 'process' evaluation referred to above.

In line with a realist understanding, the TFP (and other social policies) can be conceived of as being capable of doing certain things or as having 'tendencies' to work in particular ways but their effects (outcomes) will be contingent on the context within which they operate. This speaks to a 'generative' theory of causality, which stands in contrast to a 'successionist' understanding. The latter refers to a theory of behaviour change at the heart of experimental approaches which presuppose that it is through interventions or 'treatments' (e.g. the TFP) designed to change individual behaviour that particular outcomes (e.g. crime prevention) can be achieved and measured (Mason and Prior, 2008). A realist position recognises that social programmes are interventions in, and attempts to exert control within, the *social* world; they are delivered by and to active subjects, in open systems and therefore susceptible to interpretation and adaptation (Batty and Flint, 2012). Causality and causal pathways are therefore considered to be complex and dynamic, with causal power attributed to human agency, as well as the social structures in which they operate. The approach seeks therefore to identify mechanisms of change at individual, group, organisational and societal levels. To achieve this, qualitative empirical data (or respondents' accounts of their personal experiences) are taken seriously and situated within their wider social circumstances, thereby linking the individual with wider structures (Crouch and McKenzie, 2006). Although 'outcomes' are still the focus of evaluation, realist case studies incorporate a consideration of the range of outcomes generated (intended or otherwise), including those identified from the perspectives of parents and children (Canavan *et al*, 2016).

Given this generative view of causality, attending to context is central in realist case study research; 'context' is an essential component of explanation. Realism brings with it particular ways of understanding and attending to context however. In case study research, context can sometimes be treated more or less *descriptively* rather than *analytically,* giving rise to evidence that does not amount to context-sensitive *explanation* (the 'grey box' referred to earlier) (Welch *et al*, 2011). For case studies to be properly explanatory they need to be rooted in and aligned with clear theoretical foundations and argumentation; they require conceptual depth (Clegg, 2005). Explicit attention to the utilisation, refinement and development of theory in case study research, informs and underpins realist perspectives of causality.

In arguing for the central importance of 'theory', we are referring to two different types of theory – programme theory and substantive theory (Weiss, 2000). A realist approach assumes that implementing a policy/practice equates to testing a theory about what might cause change (a theory of change) - although that theory is not always explicitly articulated. A central objective of social policy evaluations driven by realism is to make these programme theories explicit and focal to evaluation (Greenhalgh *et al*, 2015; Cornes *et al*, 2015). Although limited, there are case study examples of where evaluators have explicitly developed theories of change to understand FIS and similar types of intervention (See e.g. Cornes, 2015; NEF consulting, 2015). In addition, substantive theory provides realist evaluation with indications to the mechanisms through which programs work, the contexts within which they will work as well as their effects. Conceptualisation (of the policy 'problem' together with its complex causes, methods of intervention and indicators of 'success') is an integral part of this. This recognises that research and policy are not value free and promotes debate about the assumptions that underpin policy. In the case of the TFP,

critical assessments of the construction of 'troubled families' and more progressive, evidence-informed conceptions of family support and family intervention are needed (Canavan *et al.* 2016; Crossley 2018). The ensuing account is formulated as a middle-range theory - theorising which is positioned between universal social laws and description, and so provides something more than a logic model (Astbury and Leeuw, 2010). It is the theorising potential of case study research that has the potential to add real value to the evidence base around FIS.

Small-scale research that provides strong contextualised explanation and is clearly embedded within extant theoretical knowledge is not stand-alone and definitive but rather 'conceptually generative' (Crouch and McKenzie, 2006). The expectation is that meaningful theoretical conceptualisation can be followed up and built upon in academic colleagues' own work and therefore transferred to different contexts, feeding into the collective process of knowledge accumulation. 'Generalisation' is understood not as universal therefore but a limited 'contingent generalisation' (Welch *et al*, 2011; Sanderson, 2009). As such each policy evaluation does not proceed 'from scratch'. Doing realist evaluation means building on theories that have been tested in previous social policy research and evaluation. This entails 'communal knowledge-building labour' and brings with it a collectivist attitude to knowledge but one that moves beyond simply citing others work (Westhorpe *et al*, 2011; Crouch and McKenzie, 2005). An evidence base which draws on this kind of knowledge is not simply additive. What counts as 'evidence' is that which can give us insight into the generative mechanisms and tendencies that help us understand how policies work. This is closely tied to Pawson's realist synthesis which 'assumes that the transmission of lessons occurs through a process of theory building rather than assembling empirical generalisations' (Pawson, 2002: 347).

If we consider the evidence of the success of TFP and other family interventions, priority need not be limited to programme evaluations and quantitative monitoring data. There is already a substantial body of case study work on the TFP and similar types of FIS. These have drawn attention to similar conceptions of 'good practice' and the ways of working that effect change, echoed in the TFP process evaluation too. This includes: low case-loads; whole-family approaches; relationship-based practice; generous time and flexibility; the recruitment and retention of high quality staff; and long-term commitment to families. These working practices might usefully be thought of as 'generative mechanisms' (Hoggart *et al*, 2014; Batty, 2013; Batty and Flint 2012; Boddy *et al*, 2016; Bond-Taylor 2017; Flint *et al*, 2011; Hayden and Jenkins 2014; Hoggett and Frost 2017; Sen, 2016), yet often these are not backed by good theory about how they work. There is a growing body of work that has begun to address the complexity of the TFP and theorise in more depth the ways in which these key mechanisms of change operate. Flint (2012) for instance has provided an empirical and theoretical account of 'the domestic visit' at the heart of FIP suggesting how and why this 'surveillance' has progressive as well as punitive dimensions. Bond-Taylor (2017) has suggested that 'ethics of care' provides a useful framework through which to understand and evaluate different aspects of frontline TFP practice. Parr (2016) has contributed to studies about relationship-based practice and 'the therapeutic alliance' through inductive theorising of the ways in which 'the family-worker relationship' engenders positive behaviour change within families. To bring more coherence to what is currently a fragmented body of knowledge, we need to learn from the findings of local evaluations, focusing on the core features of the TFP and the impact of these in specific contexts. We also need to engage meaningfully with a broader UK-wide and international evidence-base about promising approaches to family support and interventions.

In making a claim for realist cases studies, we appreciate realist case study research is not without its problems. Evaluations adopting this approach are likely to not only be resource intensive (Weiss, 1997) but complicated to execute effectively. On the one hand, realist case study research is particularly relevant to multi-dimensional social programmes such as TFP applied under different circumstances. On the other, because such programmes comprise many different elements and packages of practices designed to produce a range of different outcomes, by implication, they also comprise an amalgamation of theories about what mechanisms might work to effect change (Cornes *et al*, 2015; Sanderson, 2003). It is impossible to focus evaluative attention on all such aspects of a programme and so careful choices need to be made about where resources should be directed and what elements of programmes (or what potential 'causal links') should be focal to research.

There is also an argument that the language of some realist social research can feel overly technical and mechanistic e.g. the language of 'context-mechanism-outcome' configurations (Weiss, 2000., 1997). For it to be a feasible option for researchers, realist evaluation is best considered not as a very formal or standardised model of evaluation but as a way of thinking, a kind of logic within which certain core ideas need to be included in evaluation designs. This is also not a disinterested endeavour but rather a *critical* one; an intellectual effort focussed on uncovering implicit value assumptions underpinning social policy and an intention to drive action to change and improve social policy, according to key ethical principles (Eurochild, 2015). This can incorporate other existing and complementary evaluation approaches however. There are examples, of integrating realist ideas with action research approaches and Qualitative Comparative Analysis (QCA) for instance (Houston, 2010; Befani et al, 2007). It may also be the case that realistic evaluation is most compatible

with evidence-*informed* practice or the practice-based evidence paradigm which examines interventions within routine practice drawing on practical and community experience (Eurochild, 2015; Westhorpe *et al*, 2011; Sanderson, 2009; Webb, 2001).

These realist insights can shift current thinking towards more sophisticated social policy evaluation research designs. A new model for evaluation is required that accommodates alternative methods with realist case study research as a fundamental component. Such a model would reject the evidence hierarchy and seek to adopt a versatile, mixed–method research approach and include multiple, linked case studies of practice. While qualitative methods will have a crucial role to play with their distinct and indispensable strengths in case study research, it is important to recognise that quantitative techniques will also be important according to the specific questions being posed and the outcomes being evaluated (Canavan *et al*,2016). The latter may be used in parallel or sequentially within evaluations. Quantitative methods can be employed to develop reliable descriptions and provide accurate comparative data (McEvoy et al, 2006). They can also both assist in theory development. While quantitative measures do not in themselves establish causation, statistical associations help examine causal relationships (Bramley and Fitzpatrick, 2018). Mixed methods approaches are of course not new but realist evaluation provides a coherent epistemological rationale for their use. Beyond the TFP, several developments in evaluation research take forward these approaches. The work of the Centre for Evaluating Complexity across the Nexus (CECAN) seeks to improve policy evaluation in ways that are in accordance with the complexity of causation. A realist orientation is also being developed by the What Works Centre for Children's Social Care. In its initial work, the Centre, has sought to adapt Johnson *et al's* (2014) realist approach to reviewing and developing the evidence-base for children's social care services.

Conclusion

After over 20 years of research into family interventions, what collectively the evidence base tells us is uncertain. The first TFP national evaluation was not designed to address this knowledge vacuum and the uncertainties LAs face in deciphering what works, for whom, why and in what context. While not denying that the TFP fails to address the structural causes that underpin the many problems families face (Hargreaves *et al*, 2018), disadvantaged families still remain likely to benefit from short, medium and long-term support to improve their well-being. Ideological critique should not therefore lead to an abandonment of family support nor necessarily the TFP (Axford and Morpeth, 2013). Rather, there is a need to provide learning for TFP practitioners. There is a need to challenge the epistemological and methodological underpinnings to the TFP (and the ETFP) outcomes framework as well as to commission a broader range of, and more complex approach to, evaluation study designs in order to pose more insightful research questions and generate more extensive and inclusive evaluation findings. Case study research founded on realist principles offers a way forward and can potentially be instructive for policy development, particularly for policies and practices such as the TFP that are multidimensional, complex and develop along context-specific trajectories.

References

Astbury, B., & Leeuw, F. L. (2010). Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation. *American Journal of Evaluation*, *31*(3), 363–381. DOI: 10.1177/1098214010371972.

Axford, N and Morpeth, L (2013). Evidence-based programs in children's services: A critical appraisal. *Children and Youth Services Review*, 35(2), 268-277. https://doi.org/10.1016/j.childyouth.2012.10.017

Bailey, N. (2012) *Policy Based on Unethical Research*, http://www.poverty.ac.uk/news-and-views/articles/policy-built-unethical-research

Ball E., Batty, E. and Flint, J (2016). Intensive family intervention and the problem figuration of 'troubled families'. *Social Policy and Society*, 15(2), 263–74. https://doi.org/10.1017/S1474746415000469.

Bate, A and Bellis, A (2018). The Troubled Families Programme (England): Briefing Paper, London: Housing of Commons

Batty, E. (2013). Learning and soft outcomes: evidence from intensive intervention projects. *Social Policy and Society*, 13(3), 357–69. https://doi.org/10.1017/S1474746413000286

Batty, E and Flint, J (2012). Conceptualising the contexts, mechanisms and outcomes of intensive family intervention. *Social Policy and Society*, 11(3), 345-358. https://doi.org/10.1017/S1474746412000073

Befani, B., Ledermann, S., & Sager, F. (2007). Realistic Evaluation and QCA: Conceptual Parallels and an Empirical Application, *Evaluation*, 13(2), 171–192. https://doi.org/10.1177/1356389007075222

Bewley, B., George, A., and Rienzo, C. (2016) The Impact of the Troubled Families Programme: Findings from the Analysis of National Administrative Data, London: Department for Communities and Local Government.

Boaz, A., Grayson, L., Levitt, R and Solesbury, W. (2008) 'Does evidence-based policy work? Learning from the UK experience', *Evidence and Policy*, 4(2): 233-53

Blades, R., Erskine, C., and Day, L. (2016). National Evaluation of the Troubled Families Programme. Families' Experiences and Outcomes, London: Department for Communities and Local Government.

Boddy, J., Statham, J., Warwick, I., Hollingworth, K., & Spencer, G. (2016). What Kind of Trouble? Meeting the Health Needs of 'Troubled Families' through Intensive Family Support. *Social Policy and Society, 15*(2), 275-288. https://doi.org/10.1017/S1474746415000494

Bond-Taylor, S. (2017) 'Dimensions of family empowerment in work with so-called "Troubled Families"', *Social Policy and Society*, 14, 3, 371–84.

Bonell, C., McKee, M and Fletcher, A (2016). Troubled Families, Troubled Policy Making. The BMJ, doi: 10.1136/bmj.i5879.

Bramley, G. and Fitzpatrick, S. (2018). Homelessness in the UK: who is most at risk? *Housing Studies*, 33 (1), 96-116. https://doi.org/10.1080/02673037.2017.1344957

Butler, P. (2016). Ministers' claims about troubled families scheme 'grandiose', MPs told, *The Guardian*, 19th October.

https://www.theguardian.com/society/2016/oct/19/troubled-families-scheme-ministers-claims-grandiose-mps-told

Canavan, J., Pinkerton, J. and Dolan, P (2016). *Understanding Family Support: Policy Practice and Theory*. London and Philadelphia: Jessica Kingsley Publishers

Cairney, P (2018). The UK government's imaginative use of evidence to make policy. *British Politics*, https://doi.org/10.1057/s41293-017-0068-2

Cameron, D. (2011) 'Speech on the fightback after the riots', 15th August http://www.number10.gov.uk/news/pms-speech-on-the-fightback-after-the-riots/ accessed XX

Chen, H.T. (1990). *Theory-Driven Evaluation*. London: SAGE.

Cornes, M., Whiteford, M. and Manthorpe, J (2015) Fulfilling Lives Islington and Camden: A realist review of programme theory, London: KCL.

Clegg, S (2005). Evidence-based practice in educational research: a critical realist critique of systematic review. *British Journal of Sociology of Education*, 26(3), 415-428, DOI: 10.1080/01425690500128932

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

Crouch, M and McKenzie, H (2006). The logic of small samples in interview-based qualitative research, *Social Science Information*, 45 (4), 483 - 499. https://doi.org/10.1177/0539018406069584

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

Department for Communities and Local Government (2014). Understanding Troubled Families', London: Department for Communities and Local Government

Department for Communities and Local Government (2012a). Working with Troubled Families: A Guide to Evidence and Good Practice, London: Department for Communities and Local Government.

Department for Communities and Local Government (DCLG) (2012b) 'Listening to Troubled Families', London: Department for Communities and Local Government.

Dillane, J., Hill, M., Bannister, J. and Scott, S. (2001). *Evaluation of the Dundee Families Project.* Edinburgh: Scottish Executive/ Dundee City Council/NCH Action for Children.

Dixon, J., Schneider, V., Lloyd, C., Reeves, A., White, C., Tomaszewski, W., Green, R. and Ireland, E. (2010). *Monitoring and evaluation of family interventions (information on families supported to March 2010.* London: Department for Education.

Durose, C., Needham, C., Mangan, C., Rees, J (2017). Generating 'Good Enough' Evidence for Co-Production, *Evidence and Policy: A Journal of Research, Debate and Practice*, 13(1), 135-151.

Fletcher, A., Gardner, F., McKee, M and Bonell, C (2012) 'The British government's Troubled Families Programme.' The BMJ, *BMJ* 2012 345 e3403.

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

Greenhalgh, T., Wong, G., Jagosh, J., Greenhalgh, J., Manzano, A., Westhorp, G., and Pawson, R. (2015). Protocol – the RAMESES II study: developing guidance and reporting standards for realist evaluation. *BMJ Open,* doi: 10.1136/bmjopen-2015-008567

Gregg, D (2010). Family intervention projects: a classic case of policy-based evidence, London: Centre for Crime and Justice Studies.

Eurochild (2015). A pluralist guide to appropriate study designs/evidence types, Brussels: Eurochild

Hargreaves, C., Hodgson, P., Mohamed, J. N., & Nunn, A. (2018). Contingent coping? Renegotiating 'fast' disciplinary social policy at street level: Implementing the UK Troubled Families Programme. *Critical Social Policy*. https://doi.org/10.1177/0261018318780094

Hayden, C. and Jenkins, C (2014). Troubled Families Programme in England: "wicked problems" and policy-based evidence. *Policy Studies*, 35(2), 631-49. https://doi.org/10.1080/01442872.2014.971732

Haynes, L., Service, O., Goldacre, B and Torgerson, D (2012). Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials, London: Cabinet Office.

Hoggett, J., Ahmad, Y., Frost, E., Kimberlee, R., McCartan, K. and Solle, J. (2014) *The troubled families programme: A process, impact and social return on investment analysis.* University of the West of England, UWE Repository.

Hoggett, J. and Frost, E (2017). The troubled families programme, and the problems of success. *Social Policy and Society*, 17(4), 523-534. https://doi.org/10.1017/S1474746417000148

Houston, S (2010). Prising Open the Black Box: Critical Realism, Action Research and Social Work. *Qualitative Social Work*, 9(1), 73-91. https://doi.org/10.1177/1473325009355622

Johnson, S.D., Tilley, N. & Bowers, K.J. J(2015). Introducing EMMIE: an evidence rating scale to encourage mixed-method crime prevention synthesis reviews, *Journal of Experimental Criminology*, 11(3): 459–473. https://doi.org/10.1007/s11292-015-9238-7

Jones, A., Pleace, N., Quilgars, D. and Sanderson, D. (2006). *Addressing Anti-social Behaviour: An independent evaluation of Shelter Inclusion Project*. London: Shelter.

Marchal, B., Belle, S. van, Olmen, J. van, Hoerée, T., ans Kegels, G. (2012). Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. *Evaluation*, *18*(2), 192–212. https://doi.org/10.1177/1356389012442444

Mason, P and Prior, D (2005). The Children's Fund and the Prevention of Crime and Anti-Social Behaviour. *Criminology & Criminal Justice,* 8(3), 279–296. https://doi.org/10.1177/1748895808092430

McConnell, A. (2010). Policy Success, Policy Failure and Grey Areas In-Between. *Journal of Public Policy, 30*(3), 345-362. doi:10.1017/S0143814X10000152

McEvoy, P and Richards, D (2006). A critical realist rationale for using a combination of quantitative and qualitative methods. *Journal of Research in Nursing,* 11(1), 66–78. https://doi.org/10.1177/1744987106060192

MHCLG (2018). Supporting Disadvantaged Families: Annual Report of the Troubled Families Programme 2017-18, London: MHCLG.

National Audit Office (2016). The Troubled Families Programme: Update, London: National Audit Office.

NEF Consulting (2015).   Evaluation of Shelter's Hackney Family Service, London: NEF Consulting.

NIESR (2016). *NIESR Press Release – No evidence Troubled Families Programme had any significant impact on key objectives, NIESR evaluation finds* http://www.niesr.ac.uk/media/niesr-press-release-%E2%80%93-no-evidence-troubled-families-programme-had-any-significant-impact-key

Nixon, J., Parr, S., Hunter, C., Myers, S., Sanderson, D. and Whittle, S. (2006). *Anti-social Behaviour Intensive Family Support Projects: An evaluation of six pioneering projects*. London: Communities and Local Government.

Nixon, J., Parr, S., Hunter, C., Myers, S., Sanderson, D. and Whittle, S. (2008). *The longer term outcomes for families who had worked with Intensive Family Support Projects*. London: Communities and Local Government.

Parr, S. (2016). Conceptualising "The Relationship" in intensive key worker support as a Therapeutic Medium. *Journal of Social Work Practice*, 30(1): 1–16.

Pawson, R. (2002). Evidence-based Policy: The Promise of `Realist Synthesis'. *Evaluation*, *8*(3), 340–358. https://doi.org/10.1177/135638902401462448

Pawson, R. and Tilley, N. (1997). *Realist Evaluation*. London: Sage Publications Ltd.

Purdon, S., and Bryson, C. (2016) National Evaluation of the Troubled Families Programme. Technical report: impact evaluation using survey data. London: Department for Communities and Local Government.

Ragin, C. (2014) *The Comparative Method*, 2nd edition, Oakland: University of California Press.

Sanderson, I (2009). Intelligent Policy Making for a Complex World: Pragmatism, Evidence and Learning. *Political Studies*, 57: 699-719. doi: 10.1111/j.1467-9248.2009.00791.x

Sanderson, I (2002). Making Sense of 'What Works': Evidence Based Policy Making as Instrumental Rationality. Public Policy and Administration, 17(3), 61- 75. https://doi.org/10.1177/095207670201700305.

Sanderson, I (2003). Is it 'What Works' That Matters? Evaluation and evidence-based policy-making. *Research Papers in Education*, 18(4), 331-345. https://doi.org/10.1080/0267152032000176846

Sayer, A (2017). Responding to the Troubled Families Programme: Framing the Injuries of Inequality. *Social Policy and Society*, 16(1), 155–164. https://doi.org/10.1017/S1474746416000373

Sen, R. (2016). Building Relationships in a Cold Climate: A Case Study of Family Engagement Within an "Edge of Care" Family Support Service', *Social Policy and Society*, 15(2), 1–14. https://doi.org/10.1017/S1474746415000615

Skeggs, B (2007). *Formations of Class and Gender: Becoming Respectable*, London: Sage.

Squires, P and Measer, L (2005). 'Below Decks on the Youth Justice Flagship: The Politics of Evaluation', in Taylor, D and Balloch, S (eds), The Politics of Evaliation: Participation and Policy Implementation, Bristol: The Policy Press, pp. 21-40.

Tunstill, J., Allnock, D., Akhurst, S. and Garbers, C. (2005). Sure Start local programmes: implications of case study data from the National Evaluation of Sure Start. *Children and Society*, 19(2), 158–71

Weiss, C. H. (1997). How Can Theory-Based Evaluation Make Greater Headway? *Evaluation Review*, *21*(4), 501–524. https://doi.org/10.1177/0193841X9702100405

Weiss, C. (2000). 'Which Links in Which Theories Shall We Evaluate?', in P. J. Rogers, T. Hacsi, A. Petrosino and T. A. Huebner (eds) *Program Theory in Evaluation: Challenges and Opportunities*, *New Directions for Evaluation*, 87. San Francisco, CA: Jossey-Bass.

Webb, S (2001). Considerations on the validity of evidence-based practice in Social Work**.** *British Journal of Social Work*, 31(1), 57-79. https://doi.org/10.1093/bjsw/31.1.57.

Welch, C., Piekkari, R., Plakoyiannaki, E. et al. J (2011). Theorising from case studies: Towards a pluralist future for international business research. *Journal of International Business Studies*, 42 (5), 740–762. https://doi.org/10.1057/jibs.2010.55

Wells, P (2018). Evidence based policy making in an age of austerity. People, place and policy online, 11 (3), 175-183.

Westhorpe, G., Prins, E., Kusters, C., Hultink, M., Guijt, I and Brouwers, J (2011). *Realist Evaluation: an overview*. Seminar Report, Wageningen: Centre for Development Innovation.

White, C., and Day, L. (2016) National Evaluation of the Troubled Families Programme. Final Report on the Process Evaluation. London: Department for Communities and Local Government

Yin, RK (2014). *Case Study Research: Design and Methods* (5th edn.). Thousand Oaks: SAGE