

Chapter forthcoming in: Hodgson D, Fred M, Bailey S, Hall P, eds. *The projectification of the public sector*. Routledge, 2019

Final, 22 August 2018

Policy pilots as public sector projects: Projectification of policy and research

Stefanie Ettelt and Nicholas Mays

Abstract

Policy pilots are a cornerstone of evidence-based policy and they are frequently used to test new policies in health and social care in England. This chapter argues that policy pilots share many of the characteristics of other public sector projects and can therefore be understood as part of the phenomenon described as ‘projectification’ in public administration. Policy pilots are used to reduce the uncertainties associated with new policy, but they also allow central government to manage the relationship with local organisations such as local authorities or organisations of the NHS. Piloting also appeals to policy makers as pilots promise to promote practical and innovative solutions to often complex and contested policy problems, such as whole system transformation. Pilots attract criticism of depoliticisation, although in practice there is little suggestion that they provide the definite answers that policy-makers desire or reduce conflict over policy aims.

Keywords

Piloting, public sector projects, projectification, health and social care policy, whole system transformation, depoliticisation

Policy pilots are an intellectual cornerstone of evidence-based policy (EBP) as they hold out the promise, at least ostensibly, that policy will be rigorously tested before being rolled out more widely. Pilots are thus usually accompanied by evaluations designed to establish whether the policy change being tried out has the potential to generate the intended effects and is worth the investment. In England, national policy pilots have been promoted as an important instrument for central government, in particular, to better inform policy formulation and to manage the risks arising from new policy. While much attention has been given to the role of evaluation in generating the evidence from pilots (Cameron et al., 2011, Craig et al., 2008, HM Treasury, 2011), policy pilots have yet to be analysed as a specific type of policy project.

There is now a substantial literature analysing, and critiquing, the role of projects and project management in the public sector. Projects have become the dominant mode of organising activity in the corporate world and, in countries such as England, are widely used in the public sector. By now, some sectors are entirely organised through projects, notably international development. The European Union exercises a large swathe of its policy influence through project funding, either in regional development or through funding of the arts and sciences (see the Chapters by Fred and Mukhtar-Landgren and by Jalocha and Ćwikła in this book). Nowadays, scientific research is almost exclusively organised through competitively acquired funding in the form of project grants and commissioned research. In innovation policy of all kinds, projects are ubiquitous, with policy-making itself also becoming increasingly 'projectified' (see Hall's Chapter in this book).

Policy pilots share many of the characteristics of projects discussed in the project and project management literature. Like other public sector projects, pilots promise to accelerate change (Jensen et al., 2013), and to combine "controllability and adventure" (Sahlin-Andersson and

Söderholm, 2002), three qualities that are highly desirable in policy-making. They offer the prospect of enabling policy makers to control the risks of new policy while also creating a space for policy experimentation. They also operate within a set time frame, usually compatible with electoral cycles, occupy a particular space as a “temporary organisation” (Lundin and Söderholm, 1995) within a permanent (host) organisation, and are normally funded through distinct budgets rather than mainstream funding. Thus there are boundaries drawn between the pilot as a project and the routine work of the host organisation, aimed at focusing energy on achieving project goals by prioritising specific activities aimed at facilitating a desired change from usual practice. Policy pilots also seem to share some of the problems attributed to projects and projectification, including concerns about short-termism, the difficulty of sustaining change beyond the duration of the pilot, and the potential for depoliticisation (Parsons, 2002).

This chapter aims to contribute to the literature on projects and projectification by considering the role of policy pilots in policy-making as a form of public sector project. Our observations are derived from our research on policy piloting and our experience as evaluators of policy pilots in health and social (long-term) care in England (Ettelt et al., 2015a, Ettelt et al., 2015b). The first part examines the conceptual overlap of projects and pilots, and identifies the project characteristics of policy piloting, drawing on the literature on project and project management. In doing so, we draw attention to the dual structure of pilot programmes since they typically comprise one or more implementation projects (i.e. pilots implemented in pilot ‘sites’) and a linked research project (i.e. the evaluation of these pilots). In the second part of the chapter, we make three observations based on our research on policy piloting in health and social care in England that resonate with the literature on projects and projectification. These are: (1) the relevance of the centre-periphery relationship as the

governance context in which policy piloting takes place; (2) the importance given to evaluation and generating evidence; and (3) the difficulty of sustaining achievements and ‘scaling up’ efforts after the end of the pilots that explains their limited contribution to policy change. We conclude by reflecting on the political implications of policy piloting in health and social care in England, arguing that piloting is politically convenient as it shifts the responsibility for potential failure to local actors while reflecting positively on the aspirations of Central Government.

Policy pilots as projects

Policy piloting is a type of policy project. Pilots have a deadline, budgets, a project management team and a declared goal of promoting policy change. In instrumental terms, they are a change management tool that policy-makers can use when, or before, introducing policy change more widely. In English policy-making, pilots typically come in clusters, as pilot programmes, which then consist of a number of projects that are embedded and layered in a pilot programme, and are given a distinctive name such as ‘demonstrators, ‘pioneers’ or ‘trailblazers’. We here distinguish three layers within a national policy pilot programme: a national policy project that is the pilot programme; pilot projects through which the policy is implemented (and tested) locally; and a research project that aims to find out whether the policy ‘works’ by evaluating the local policy pilots (occasionally, this includes evaluating the effectiveness of national level management of the programme). In this model, drawing on Jensen and colleagues (in this book), policy piloting often combines characteristics of the “change project”, intended to change existing organisational practices, with those of the “experimental project”, aimed at testing new activities to be added to an existing practice.

In England, central Government initiates and funds policy pilots with the stated aim to test policy before it is rolled out more widely.¹ Such programmes typically consist of several local pilots, but the number of sites per programme varies widely. These are implemented, for example, by local authorities, local organisations within the National Health Service (NHS), or other organisations in the public, private or voluntary sector, often in combination (e.g. local authorities working with NHS Clinical Commissioning Groups and care homes that are owned by private businesses or voluntary organisations). National policy-makers, in this case usually officials at the Department of Health and Social Care or one of its arm's length agencies such as Public Health England or NHS England, also commission an evaluation, typically involving a research team, often, but not exclusively, based in academia.

Lundin and Söderholm (1995) distinguish 'task', 'time', 'teams' and 'transition' as central dimensions of projects and we will use these attributes to further disentangle the organisation of health and social care policy pilots. The *task* of pilot sites is to implement the policy that is being piloted, meaning that their task is to operationalise policy and put ideas set out in policy papers into organisational practice. In this sense, implementation in pilot sites can be seen as a "single project as unit of analysis" that is "understood as a manageable and researchable item whose intrinsic mechanisms were to be uncovered in pursuit of project success" (Packendorff, 1995, Packendorff and Lindgren, 2014). Project success, however, is typically not clearly specified *ex ante* and pilot sites can vary hugely in how they interpret national policy aims and what they aim to achieve. Instead, it is often the task of the evaluators to define clearly the outcomes sought and the related indicators that help specify and measure these outcomes, and it is their responsibility then to establish whether these outcomes can be attributed to the policy that is being piloted.

¹ We refer to England instead of the United Kingdom (UK), as responsibility for health and social care policy is devolved to the four countries that form the UK.

The task of implementing policy in pilot sites is usually presented as orderly and straightforward, but inevitably turns out to be as messy and complex as implementation in any other policy ‘swamp’ (Argyris and Schön, 1978). Pilot implementation includes a large number of decisions, trial and error, and eliminating alternatives. It also involves ‘unlearning’ previous ways of working and typically relies on actors within the host organisation and its wider network of public and private sector organisations to come together and agree on a course of action, especially in those pilots that aim for ‘integration’ between sectors. Policy implementation is complicated by the fact that it is sometimes not clear what is to be implemented. A broad policy idea may appear plausible in a Green Paper in which a government consults on its plans for future policy, but remains aspirational if there is no clear idea of how it can be operationalised in practice.

The research project is also broken down into a number of sub-projects, usually an outcome evaluation, process evaluation and an economic evaluation, to establish whether the policy can produce certain outcomes, how these are produced, and whether it constitutes value for money. The task therefore is to complete the evaluation and produce findings within given timeframes and budgets. Compared to the task of implementing the pilots, the research project is significantly more structured *ex ante* with a predefined sequence of tasks (e.g. research ethics and governance approvals, development of tools, data collection, data analysis, etc.) set out in a proposal or protocol. There is space for adjustments, but this depends on the research design with some designs being more flexible than others (e.g. randomised controlled trials rely on a stricter protocol and a more narrow definition of the nature of the policy as an ‘intervention’ than, say, case study research).

It is interesting to consider what constitutes the ‘task’ from the perspective of national policy-makers who initiate pilots and commission their evaluations. Their task usually includes

clarifying policy objectives, selecting pilot sites and developing implementation guidelines. However, as our research has shown, objectives of pilots are rarely stable and are likely to change over time, especially in an environment in which policy-makers change frequently, with Ministers being replaced and civil servants reorganised or made redundant to reduce the Government's administrative headcount (Ettelt et al., 2015a). After the end of a pilot programme, the tenets of evidence-based policy assert that policy-makers 'make use' of the findings from the evaluations and deliberate on their implications for the roll-out of the policy. In some cases, evaluation findings are instrumental in informing policy decisions. Yet, in other cases, priorities for policies have shifted during the life of the programme. If findings are less favourable than expected, policy-makers may be inclined to drop the policy, postpone its wider implementation, or try to find solutions to overcome problems identified by the research. Alternatively, they could 'cherry pick' findings to make them sound more positive in pursuit of wider policy goals (e.g. in the case of the evaluation of the Whole Systems Demonstrators², which were testing the use of assistive technologies for people with long-term conditions, the government highlighted the positive findings in order to justify further technological investment) (Ettelt et al., 2015b).

Time is of obvious relevance to policy pilots, which are time-limited by definition. Since it is the declared purpose of policy pilots to inform policy decisions, the programme needs to be organised in a way to allow for the evaluation to report when the findings are needed. Thus, pilot programmes often depend on political schedules (e.g. legislative cycles) to determine their endpoint. The problem is that this can leave too little time to implement the pilots, let

² The 'Whole System Demonstrators' were a government-funded pilot programme conducted between 2008 and 2011 that tested the use of assistive technologies such as pendant alarms for older people at risk of falling (referred to as telecare) or diabetes management devices (referred to as telehealth) used in people's homes. The name of the programme was aspirational, indicating the intention to use these technologies to help integrate services at the boundary between the health and social care systems. While the evaluation of the individual technologies showed moderate positive effects, the programme did not achieve its overarching aim of integrating systems of care.

alone to evaluate them, especially if the pilots need to produce a sufficient number of observable outcomes. There is an increasing recognition by policy-makers that more time needs to be allocated to allow pilot programmes to be set up. However, it is often hard to predict how much time is sufficient for pilot sites to operationalise the policy and when a programme is mature enough for its evaluation to capture the policy's true potential (Ogilvie et al., 2011). It is sometimes claimed that policy pilots would have produced better outcomes if they had been allowed more time and that this explains why very similar pilot programmes are frequently mounted in swift succession (Bardsley et al., 2013). Yet while it seems obvious that programmes should be allowed to 'bed in', this is often not realistic within the time constraints that come with national policy-making. Lurking behind this empirical problem is an unresolved question as to how much time is enough to be able to judge whether a programme has generated sufficient evidence to determine whether the policy is worth persisting with or should be discontinued.

The pilot projects and their research projects are typically executed by separate *teams*. Staff implementing the pilots may be recruited from the permanent workforce of the organisation or employed on temporary contracts, with pilot managers often having other responsibilities within the host organisation. The project literature emphasises that the project team simultaneously anchors the project in the structure of the host organisation and separates it from its routine operations (Lundin and Söderholm, 1995, Jensen et al., 2013, Godenhjelm et al., 2015). This resonates with pilot projects that typically rely on project managers to collaborate extensively with other members of the organisation and with other organisations, while at the same time keeping the team focused on implementing change. In addition, many contemporary pilot programmes in England aim to improve the collaboration between different organisations (e.g. to promote 'integration' of health and social care). In such programmes, pilot teams can sit between existing organisations, making it more difficult to

maintain links with the existing structures and secure their support. Typically, such pilots do not sit within a legal entity, nor do they form their own one, which means they are not allowed to hold their own budget, which arguably exposes them even more to tensions between, or simply lack of interest from their participant organisations.

The research project is also organised around a team, typically brought together for the purpose of the evaluation and based on individual competencies. Temporary contracts are the norm for the majority of research staff with the exception of senior researchers who tend to work on other projects alongside contributing to a particular evaluation. There are different schools of thought with regard to the separation of implementers and evaluators (more about this below), yet evaluators are typically expected to be independent from policy-makers, for example, by limiting the ability of policy-makers to influence, or prevent, the publication of findings, although this varies across different policy fields with health being the most wedded to independent evaluation.

The stated purpose of national piloting policy is to promote policy change and thus facilitate **transition**. While this is in line with other types of public sector projects (Jensen et al., 2013), policy piloting, at least in theory, aspires to a different route to transition: policy change based on evidence produced from evaluation. The assumption is that findings from evaluation will be used by policy-makers to inform future policy decisions, so that these decisions are taken on a more informed basis, which then leads to decreased opposition and increased support for the policy change. However, this instrumental logic is problematic. Findings appear to be most palatable when they confirm a decision, rather than challenge it. They are also not always conclusive enough to support a specific decision. Frequently, more research is needed, for example, because the pilot sites have taken longer than expected to implement the intervention and more time is needed to measure patient/client outcomes in comparison

with some sort of 'control' group. In most cases, it is difficult to generalise beyond the remit of the programme and questions arise about the transferability of findings to other prospective areas.

More familiar from a project perspective is the idea that pilots are used to 'pump prime' change nationally by initiating projects locally from which other places can then learn. This intention is reflected in programmes in which pilot sites are selected based on their experience with, and success in, implementing existing policy. However, this approach generates concerns about a potential lack of sustainability of local change and the difficulty of 'rolling out' the programme to less 'expert' sites, especially if pilot sites have been selected primarily to be able to 'show off' their achievements (Ettelt et al., 2015a). There can also be substantial confusion about the mechanism of change that policy pilots are expected to promote, which in turn can have an impact on the options and opportunities for evaluation. For example, if pilot sites are selected because of their experience, it will be difficult to organise a trial that measures the likely costs and benefits when sites start implementing 'from scratch' (Hendy et al., 2012).

To summarise, policy pilots share many of the characteristics of other types of public sector projects, but there are also some qualifications. The task of implementing policy pilots is often less clear in operational terms than it appears in policy documents. Pilots also sit within a political environment that both enables and compromises the tasks of implementation and evaluation. As pilots are a temporary organisation, there is usually a firm deadline to work towards, but deadlines are typically politically determined, not based on a realistic estimate of the time it is likely to take to implement the policy change and have it 'bed down' sufficiently to allow for meaningful evaluation of its long-term consequences. The different teams involved emphasise the separation of implementation, evaluation and policy-making, and also

highlight that policy piloting is a multiple team effort that requires substantial coordination. Finally, while facilitating change is the obvious purpose of policy piloting, there can be confusion as to how transition is to be brought about and the role of evaluation in facilitating policy change.

Three observations from policy piloting in health and social care in England

In the next section, we make three observations based on the findings from our research on nationally initiated policy pilots in health and social care in England that resonate with debates in the projectification literature. These are 1. the observation that national policy piloting is embedded in, and relies on, the specific centre-periphery relationships prevalent in health and social care governance in England; 2. the importance of generating evidence from evaluation in policy piloting as its proclaimed mechanism of change; and 3. the difficulty of sustaining achievements in policy implementation and rolling them out following the end of a pilot programme.

1 – Piloting cuts across the centre-periphery relationship

National policy piloting relies on local actors to implement national policy and therefore spans the centre and periphery involved in policy-making. In England, central Government initiates national policy pilots by inviting local actors to participate in a pilot programme, sometimes, but not always, associated with the promise of additional funding. Local statutory actors who volunteer to participate in the programme then implement the pilots in their locality, often in collaboration with other organisations in the voluntary or private sector (e.g. care homes) or across policy sectors (e.g. NHS organisations working with elected local

authorities). In addition, there are all sorts of local and regional initiatives that may be pilot programmes, but are not usually piloting national policy (Bailey et al., 2017). The centre also commissions any national evaluation. At times, this is complemented by a requirement to invest in additional local evaluation. Central Government also tends to support networking activities between pilot sites to foster the exchange of local experience, knowledge and inspiration. These activities are typically run separately from the evaluation process. Before the financial crisis of 2007-08, there would often be a substantial support infrastructure at the centre for some pilot programmes, but this has been massively reduced in recent years, in part, in response to financial pressures, but also reflecting a new vision for a more restrained, scaled-down role for central Government in policy innovation.

Policy piloting – as practised in health and social care policy in England - reflects an understanding of an existing mode of governance and its established centre-periphery relationships. The received wisdom is that central Government is responsible for policy formulation but has limited ability to engage in, or even control implementation. There are many ways in which central influence filters through the system, but responsibility for implementation is largely delegated to local actors (Flinders, 2002, Exworthy and Frosini, 2008, Pratchett, 2004). Local actors, in turn, including local government, do not participate in national policy formulation. They may be consulted or ‘listened to’ and their staff may or may not support implementation in their role as ‘street-level bureaucrats’ (Lipsky, 1980), but they do not have a formal role in decision-making (Ling, 2002). There are of course differences between the two policy sectors: Central Government is more or less directly responsible for the NHS and overseen by its administrative agency, NHS England. Social care, in contrast, is organised entirely by elected local government. Local government, however, is dependent on a centrally allocated social care grant and operates within a centrally set policy and regulatory framework. Yet local authorities have a larger degree of

separation from the centre than, say, organisations of the NHS. They raise some of their own funds locally from business rates and household taxes and thus do not depend entirely on national resource allocation. Their leaders are primarily accountable to local voters rather than central Government, even though local councillors also have to account for their use of central government funds. Despite this, the centre acts as the initiator of policy piloting, while leaving 'implementation' to local actors. In recent years, there has been renewed interest in the idea of 'localism' and the devolution of responsibilities from the centre to the periphery (mostly to the large metropolitan centres through a series of 'devolution deals') (Lowndes and Pratchett, 2012, Cox, 2010), but on the whole Central Government in England still conceives of itself as the main driver of policy innovation.

In practice, policy ideas launched by the centre are often broad and still require a substantial amount of thought, work and effort to operationalise. The pilot is thus a mechanism to delegate the task of operationalisation to local actors who deal with the hard graft of working through the problems that implementation throws up. For example, as a policy idea, the integration of health and social care seems to make eminent sense as it promises both a better service for patients and a possibility of reducing costs in the face of increasing demands from an ageing population. Yet integration has so far been largely elusive (although it has been tried in multiple ways and many places) and consequently central Government has initiated a series of locally implemented pilot programmes to test different approaches, establish the evidence and diffuse the knowledge of how better integration can be achieved. System integration remains out of reach, but the effort is ongoing and the idea of integration, despite its many setbacks, is still very much alive. An alternative policy approach would have been to bring the health and social care systems under one legislative umbrella, develop a unified funding approach and systematically organise a (more) integrated approach to system governance and service delivery. Yet this would require national politicians and their policy

advisers to resolve a series of awkward policy dilemmas made more difficult by tightly constrained public finances. So far, this approach has been seen as 'too difficult' to be contemplated. In contrast, initiating a series of local pilots allows for an incremental approach to promoting change by devolving the responsibility for progress to local actors.

It also shifts the risk of policy failure from the centre to the periphery, while leaving the existing governance infrastructure intact. From a central Government perspective, there are at least three types of risk related to policy 'failure'. First, there is the practical risk of a policy failing either because its implications were insufficiently thought through, or because other stakeholders, on whom implementation relies, withhold their support. The project management literature emphasises the risks of time overrun and budget overspend of projects, especially in the public sector (Jensen et al., 2013). By initiating a pilot programme, Government is able to control these risks as the programme is given a fixed budget and timeline within which pilot implementers have to operate. There have been occasions in which the length of a programme has been extended, typically to give the evaluation more time to measure outcomes, but these extensions have been limited in scope. Perhaps more importantly, if the Government decides after the end of a programme that it no longer wishes to roll out the policy it is able to abandon it (although there is still the risk of some reputational damage).

Second, there is the reputational risk for Government of being seen as complacent. By initiating a pilot programme, and commissioning its evaluation, the Government gives the policy some prominence, devotes (limited) resources to it, and makes a commitment to engage with the knowledge produced from the process. Initiating pilots signals that the Government intends to address a problem and that it is willing to embark on a new approach. In this respect, national policy piloting is no different from other public projects that cater for

a demand for change by signalling innovation and entrepreneurship (Sahlin-Andersson and Söderholm, 2002) and a “crystallisation of intent” (Pellegrinelli, 2011: 236). However, Government often seems most comfortable about its pilot programmes at the stage when they are announced, suggesting that there is symbolic value in initiating pilots and commissioning their evaluation, as opposed to having to deal with their results (Ettelt et al., 2015a).

The third risk that central Government will aim to avoid is the risk of opposition to the policy growing, and of losing control over the policy discourse. Opposition to a policy is more likely to be manageable when the policy is contained within a pilot programme, which is ostensibly only meant to ‘test’ the policy on a limited scale. This gives less room for critics to oppose the policy. The management of this third type of risk through piloting contributes to concerns about the depoliticisation of policy-making. By focusing attention on local implementation and testing, piloting creates a distance between the politics of policy-making and the Government that is testing its implementation. In a similar fashion, it is much more likely that policy pilots are evaluated locally, than evaluation being applied to projects that Government embarks on within its own ranks (e.g. projects undertaken by the Cabinet Office). It also focuses on questions of policy effectiveness and cost-effectiveness, while other issues, such as distributional fairness (e.g. in respect to finding alternative ways of collecting and distributing funding for social care, which has been discussed as a matter of urgency for many years) can easily be side-lined. Thus while piloting is a mechanism for testing policy, it is also a tool to contain criticism and control the discourse, by which the centre maintains its power over policy-making, while simultaneously delegating the practical problem solving to local actors.

Many projects involve monitoring and evaluation, but in policy piloting generating evidence from evaluation is usually a key purpose of the exercise at least ostensibly and at the outset, which sets it apart from other types of policy projects (Cabinet Office, 2003, Sanderson, 2002). The idea is that by evaluating a pilot programme, insights can be gained that can inform national policy-making and create the 'evidence' in support of a policy. The hope is that the evidence generated from the pilots is sufficiently conclusive to support a decision and the typical assumption is that it will demonstrate that the new policy 'works', i.e. it will produce the effects desired.

However, in practice, the evaluation process often does not work like this. Frequently, the findings from the evaluation are equivocal. Evaluators often find it difficult to measure outcomes to an extent that they are conclusive, with some recent pilot programmes simply unable to attract the number of study participants required to measure outcomes validly in the time available. In such cases, evaluation is likely to produce a good understanding of the difficulties of implementing the pilots, but it is much harder to establish whether the mechanism underpinning the policy can generate the desired outcomes and how these outcomes may differ in a wide range of different contexts.

However, from an implementation perspective, there is also the question as to whether the evaluation is targeted at the right audience. Should the evaluation be aimed at informing mainly local policy implementers rather than central Government decision-makers? There are a number of approaches to evaluation that do exactly this, such as action-research that aims to involve implementers strongly in the research to foster learning while the pilots are being developed; but these are rarely requested when evaluations of national policy pilots in health and social care are commissioned. This also raises the question as to whether findings from evaluation should be made available more quickly (as in 'real-time' evaluation), rather than at

the end of the programme as is customary and more feasible for rigorous outcome evaluation. In practice, evaluators are often encouraged to report early and often, yet it is entirely possible that early findings differ from, even contradict, later conclusions that take account of all the data collected during a programme. Though each of these propositions sounds desirable, they tend to conflict with the idea of outcome evaluation that is both scientifically robust and independent, since this requires that the influence of implementers (and policy-makers) is kept to a minimum.

There can also be tension between the purpose of the pilots and the purpose of the evaluation, suggesting conflicting ideas about the mechanisms through which policy change influences changes in local practice. Should pilots be organised to enable evaluation to be as robust as possible? If so, local variation between interventions should presumably be kept to a minimum. Or should the pilots be more flexible and allow more variation to encourage local innovation, in which case the chances of robust outcome evaluation would be reduced dramatically? Having substantial variation between different pilots within a programme means that effects of the policy cannot be pooled across pilot sites (as they all implement something different), hence limiting the 'evaluability' of the programme (Ogilvie et al., 2011). Presumably, such decisions, if taken consciously, hinge on different ideas about the mechanisms of policy change. From an evidence-based policy perspective, it is assumed that policy change should be determined by the best available evidence; robust evaluation is therefore vital. Yet policy piloting also allows for a more bottom-up type of policy change that invites local actors to experiment less formally and develop their own solutions to policy problems. This mechanism of policy change tends to be favoured by local organisations, perhaps especially local authorities; yet it can conflict with ideas of evidence use in policy that favour robust evaluation and a more prescriptive approach to policy piloting.

3 – The difficulty of sustaining achievements and spreading them ('scaling up')

To allow project managers in pilot sites to develop new ways of working, the temporary organisation that is the pilot requires a degree of separation from the permanent (host) organisation. At the same time, project managers typically rely on the host organisation, its resources, structures and networks, for support and to achieve the effects they are expected to produce.

Yet by separating the pilots from their host organisations, the pilots face the same challenge that many other projects face: how can the achievements of the pilot be 'mainstreamed' into the normal ways of working of the host organisation? For example, if the pilots set out to test the impacts of providing frail individuals with telecare devices to improve their care at home (as attempted in the Whole System Demonstrators), how are local authorities expected to continue offering these devices and provide the new service alongside existing services after the funding for the pilots has expired? This problem is not trivial, as few policy pilot programmes appear to have been sustained beyond the end of the programme (though this may be an artefact of the fact that this phenomenon is difficult to research). Those interventions that have continued after the formal end of a pilot programme have often been scaled down. Given current funding constraints, at least in England, it is difficult to imagine how this could be different, unless the intervention piloted clearly shows it can save money. Arguably, this is neither unexpected, nor necessarily undesirable. It is not unexpected, given that much research has demonstrated that it takes a long time to embed almost any type of change into routine practice; and it may not be undesirable, if we concede that piloting should be undertaken without fixed expectations as to its results, similar to the concept of equipoise that justifies randomised clinical trials (Petticrew et al., 2013).

There is an added challenge for policy piloting, arising from the aspiration associated with evidence-based policy. How can the achievements of the pilots be ‘scaled’ and rolled out nationally? If the intention is that after a successful experience (however defined), the policy is to be implemented in other areas beyond those that participated in the pilot programme, how can the learning from the pilots inform these efforts? Evidence-based policy stipulates that the evaluation will generate the generalisable knowledge that will help others to follow suit. However, evaluation, especially of the national, outcome-focused variety, is often not well placed to provide the level of detailed, contextual knowledge that implementers are likely to require if they are to extend the innovation to other places. In addition, those studies with strongest claims to internal validity (e.g. research designs that test causality such as randomised controlled trials) tend to have particularly limited external validity, i.e. they cannot easily be generalised or their findings applied to other places, and researchers may hesitate to provide the kind of prescriptive steer desired by project managers and policy makers.

One aspect of this problem is the relationship between the pilot, and the host organisation and its wider network. In the language of piloting and evaluation, this is how the ‘intervention’ relates to the ‘context’ or, more precisely, how the variables associated with the context influence the intervention/policy and thus the outcomes observed. One way of examining this relationship is to require pilot sites to standardise the intervention as much as possible. However, this requires prior definition of the intervention and agreement among the parties that this definition is the correct one to pilot, which is not always the case, particularly not with genuinely innovative pilots. Standardising the implementation may help the robust measurement of outcomes, taking account of contextual variables, but it does so by precluding the option of exploring alternative approaches to operationalising the policy and

by reducing the options for adaptation that may make the experience of participating in the programme more meaningful to local sites. .

This problem is amplified where pilots set out to make a significant change to existing practices and systems. There can be a significant contrast between the scale of the ambition and the ability of pilot programmes to achieve it, almost irrespective of how the evaluation is conducted. Programmes such as the Whole System Demonstrators ostensibly set out to show how the ‘whole’ local health care and social care systems could become involved in the integration of services by introducing patient-operated supportive technologies (telehealth and telecare). This did not materialise during the pilot programme because too few people could be identified who had both a health care and a social care need that could be appropriately met via a technical intervention, so there was little opportunity to integrate care in this way. With hindsight, it seems curious that policy makers should believe that telecare and telehealth could be expected to transform entire local health and care systems or at least to make a noticeable difference to any of them. If this was ambitious, this was matched by the high expectations placed on the evaluation in that it was meant to provide definitive proof of the superiority of such technology compared with usual care. Not surprisingly, this did not materialise either.

Final reflection: policy piloting in the shadow of ‘politics’

In this chapter, we have argued that policy pilots have much in common with other types of public sector projects, but that there are also some differences insofar as policy piloting in our experience tends to put special emphasis on national evaluation and a proclaimed commitment to test whether pilots produce desired outcomes. By drawing on our research on policy piloting in health and social care in England, we have shown that national policy

piloting is ostensibly embedded in a particular version of public sector governance, in which central Government formulates policy and initiates national programmes while local actors solve problems of implementation. Policy piloting is therefore directly exposed to the ‘politics’ of central Government policy making, its contestation, vagaries and uncertainties, and its desire to control the risks associated with policy change, which can be practical (avoiding policy failure) as well as reputational (by delegating risk) and political (by defusing opposition).

We have also highlighted the role of evaluation in policy piloting and its importance as a mechanism of policy change, which is in line with the central tenets of evidence-based policy. However, this approach means that the stakes for evaluation are high in policy piloting and much depends on the ‘evaluability’ of the pilots as they are implemented locally (Ogilvie et al., 2011, Ettelt and Mays, 2015). Hence, policy-makers are faced with a dilemma: should pilots be organised to allow for robust evaluation and therefore risk stifling flexibility and local innovation; or should they be organised to encourage local variation, yet at the price of reducing the possibility of evaluation to assess whether the policy ‘works’ in a more generalizable sense? A further, related observation focuses on the difficulty facing policy pilots in maintaining their achievements and transferring their experience to other localities that have not participated in the programme, but are expected to learn from the pilots, especially through their evaluation. The long-term effects of policy pilots in health and social care on local change have yet to be investigated in any depth. There is much less research on long-term developments in local governments than on the short-term implementation of evidence-based practice (May et al., 2007).

Examining policy piloting through the lens of the project and projectification literature leads us to the question as to why policy-makers embark on pilot programmes. Based on the

analysis presented above, we conclude that piloting allows policy-makers to influence the policy process in at least three ways: first, policy piloting shifts the attention from the question of the desirability of a new policy to the practical problems of implementation. The pilot evaluation will tend to focus on ‘how’ the policy is interpreted and translated into practical change rather than ‘why’ it is being suggested in the first place. Second, it allows policy-makers to delegate the responsibility for success and failure of pilots and policy to local actors, and away from the centre. This is particularly tempting when the risk of failure is high, since it is difficult, both conceptually and practically, to distinguish whether ‘failure’ (in itself a socially constructed phenomenon) was the result of a conceptually flawed policy or because local actors did not implement it correctly (policy failure versus implementation failure). Third, piloting helps to mobilise claims to objectivity derived from independent, outcome-focused evaluation, which strengthens the authority of policy-makers by lending scientific authority to policy decisions (Bijker et al., 2009, Weingart, 1999). In conjunction, these strategies can partly, and perhaps only temporarily, depoliticise the policy process, an observation made by critical commentators on the claims of evidence-based policy-making, as well as on projectification in the public sector more widely (Parsons, 2002, Cicmil and Hodgson, 2006).

This is striking, given the obvious mismatch between the scale of the ambition loaded on programmes aimed at whole system change, system transformation or integration between public policy sectors and the deliberately limited investment of political capital and resource associated with most pilots. Whether this suggests the survival of pragmatism in English policy-making (the optimist’s view) or an attempt to obfuscate politics by devising a smoke-screen behind which to hide a lack of willingness to address complex social problems (the cynics’ view) remains a matter of debate (including between the authors of this chapter).

However, the ambiguity associated with these two readings may explain why policy piloting,

although initially a New Labour initiative, has remained popular with successive (Conservative) governments, including those devoted to the 'small state' and austerity governance. Projectification helps policy-makers to manage a whole set of complex relationships, uncertainties, and expectations that they would otherwise have to deal with much more explicitly and laboriously.

References

- ARGYRIS, C. & SCHÖN, D. A. 1978. *Organizational learning*, Reading (Mass.), Addison Wesley.
- BAILEY, S., CHECKLAND, K., HODGSON, D., MCBRIDGE, A., ELVEY, R., PARKIN, S., ROTHWELL, K. & PIERIDES, D. 2017. The policy work of piloting: Mobilising and managing conflict and ambiguity in the English NHS. *Social Science & Medicine*, <http://dx.doi.org/10.1016/j.socscimed.2017.02.002>.
- BARDSLEY, M., STEVENTON, A., SMITH, J. & DIXON, J. 2013. Evaluating integrated and community-based care. How do we know what works? London: Nuffield Trust.
- BIJKER, W. E., BAL, R. & HENDRIKS, R. 2009. *The paradox of scientific authority. The role of scientific advice in democracies*, Cambridge (Mass), MIT Press.
- CABINET OFFICE 2003. *Trying it out: The role of pilots' in policy-making: Report of a review of government pilots*, London, Cabinet Office, Strategy Unit.
- CAMERON, A., SALISBURY, C., LART, R., STEWART, K., PECKHAM, S., CALNAN, M., PURDY, S. & THORP, H. 2011. Policy makers' perceptions on the use of evidence from evaluations. *Evidence & Policy: a journal of research, debate and practice*, 7, 429-447.
- CICMIL, S. & HODGSON, D. 2006. Making projects critical: Introduction. *Making projects critical*. Palgrave.
- COX, E. 2010. Five foundations of real localism. Newcastle Upon Tyne: Institute for Public Policy Research/IPPR North.
- CRAIG, P., DIEPPE, P., MACINTYRE, S., MICHIE, S., NAZARETH, I. & PETTICREW, M. 2008. *Developing and evaluating complex interventions: the new Medical Research Council guidance*.
- ETTELT, S. & MAYS, N. 2015. Advice on commissioning external academic evaluations of policy pilots in health and social care: a discussion paper. London: Policy Innovation Research Unit.
- ETTELT, S., MAYS, N. & ALLEN, P. 2015a. The Multiple Purposes of Policy Piloting and Their Consequences: Three Examples from National Health and Social Care Policy in England. *Journal of Social Policy*, 44, 319-337.
- ETTELT, S., MAYS, N. & ALLEN, P. 2015b. Policy experiments: Investigating effectiveness or confirming direction? *Evaluation*, 21, 292-307.
- EXWORTHY, M. & FROSINI, F. 2008. Room for manoeuvre? Explaining local autonomy in the English National Health Service. *Health Policy*, 86, 204-212.
- FLINDERS, M. 2002. Governance in Whitehall. *Public Administration*, 80, 51-75.

- GODENHJELM, S., LUNDIN, R. A. & SJÖBLOM, S. 2015. Projectification in the public sector - the case of the European Union. *International Journal of Managing Projects in Business*, 8, 324-348.
- HENDY, J., CHRYSANTHAKI, T., BARLOW, J., KNAPP, M., ROGERS, A., SANDERS, C., BOWER, P., BOWEN, R., FITZPATRICK, R. & BARDSLEY, M. 2012. An organisational analysis of the implementation of telecare and telehealth: the whole systems demonstrator. *BMC Health Services Research*, 12, 403.
- HM TREASURY 2011. The Magenta book: Guidance for evaluation. London: HM Treasury.
- JENSEN, C., JOHANSSON, S. & LÖFSTRÖM, M. 2013. The project organization as a policy tool in implementing welfare reforms in the public sector. *International Journal of Health Planning and Management*, 28, 122-137.
- LING, T. 2002. Delivering joined-up government in the UK: Dimensions, issues and problems. *Public Administration*, 80, 615-642.
- LIPSKY, M. 1980. *Street-level Bureaucracy: Dilemmas of the Individual in Public Services*, New York, Russell Sage Foundation.
- LOWNDES, V. & PRATCHETT, L. 2012. Local governance under the coalition government: Austerity, localism and the 'Big Society'. *Local Government Studies*, 38, 21-40.
- LUNDIN, R. A. & SÖDERHOLM, A. 1995. A theory of the temporary organization. *Scandinavian Journal of Management*, 11, 437-455.
- MAY, C., FINCH, T., MAIR, F., BALLINI, L. & AL., E. 2007. Understanding the implementation of complex interventions in health care: the normalization process model. *BMC Health Services Research*, 7.
- OGILVIE, D., CUMMINS, S., PETTICREW, M., WHITE, M., JONES, A. & WHEELER, K. 2011. Assessing the evaluability of complex public health interventions: five questions for researchers, funders, and policymakers. *Milbank Quarterly*, 89, 206-225.
- PACKENDORFF, J. 1995. Inquiring into the temporary organization: new directions for project management research. *Scandinavian Journal of Management*, 11, 319-333.
- PACKENDORFF, J. & LINDGREN, M. 2014. Projectification and its consequences: Narrow and broad conceptualisations. *South African Journal of Economic and Management Sciences*, 17, 7-21.
- PARSONS, W. 2002. From muddling through to muddling up-evidence based policy making and the modernisation of British Government. *Public policy and administration*, 17, 43-60.
- PELLEGRINELLI, S. 2011. What's in a name: Project or programme? *International Journal of Project Management*, 29, 232-240.

PETTICREW, M., MCKEE, M., LOCK, K., GREEN, J. & PHILLIPS, G. 2013. In search of social equipoise. *BMJ*, 347, 18-20.

PRATCHETT, L. 2004. Local autonomy, local democracy and the 'New Localism'. *Political Studies*, 52, 358-375.

SAHLIN-ANDERSSON, K. & SÖDERHOLM, A. 2002. *Beyond project management: New perspectives on the temporary-permanent dilemma*, Stockholm, Liber.

SANDERSON, I. 2002. Evaluation, policy learning and evidence-based policy making. *Public administration*, 80, 1-22.

WEINGART, P. 1999. Scientific expertise and political accountability: paradoxes of sciences in politics. *Science and Public Policy*, 26, 151-161.

Information about the authors

Dr Stefanie Ettelt is Associate Professor in Health Policy in the Department of Health Services Research and Policy at the London School of Hygiene and Tropical Medicine. She has a background in political science. Most of her work currently focuses on policy experimentation and evaluation, policy and institutional analyses, and the use of evidence in health policy making.

Nicholas Mays has been Professor of Health Policy in the Department of Health Services Research and Policy at the London School of Hygiene and Tropical Medicine since 2003. He also directs the Department of Health and Social Care-funded Policy Innovation Research Unit. He has a background in social policy.

Contact details

Stefanie Ettelt

Department of Health Services Research and Policy

London School of Hygiene and Tropical Medicine

15-17 Tavistock Place

London WC1H 9SH

United Kingdom

Email: stefanie.ettelt@lshtm.ac.uk

Phone: +44 20 7927 2061

Nicholas Mays

Department of Health Services Research and Policy

London School of Hygiene and Tropical Medicine

15-17 Tavistock Place

London WC1H 9SH

United Kingdom

Email: nicholas.mays@lshtm.ac.uk

Phone: +44 20 7927 2222