

---

**Perspectives on Improvement and  
Effectiveness:  
Key Definitions and Concepts**

---

Adele Rowe and Christopher Chapman



**What Works Scotland (WWS)** aims to improve the way local areas in Scotland use evidence to make decisions about public service development and reform.

We are working with Community Planning Partnerships involved in the design and delivery of public services (Aberdeenshire, Fife, Glasgow and West Dunbartonshire) to:

- learn what is and what isn't working in their local area
- encourage collaborative learning with a range of local authority, business, public sector and community partners
- better understand what effective policy interventions and effective services look like
- promote the use of evidence in planning and service delivery
- help organisations get the skills and knowledge they need to use and interpret evidence
- create case studies for wider sharing and sustainability

A further nine areas are working with us to enhance learning, comparison and sharing. We will also link with international partners to effectively compare how public services are delivered here in Scotland and elsewhere. During the programme, we will scale up and share more widely with all local authority areas across Scotland.

WWS brings together the Universities of Glasgow and Edinburgh, other academics across Scotland, with partners from a range of local authorities and:

- Glasgow Centre for Population Health
- Healthcare Improvement Scotland
- Improvement Service
- Inspiring Scotland
- IRISS (Institution for Research and Innovation in Social Services)
- Joint Improvement Team
- NHS Health Scotland
- NHS Education for Scotland
- SCVO (Scottish Council for Voluntary Organisations)

This Literature Review is one of a series of papers that What Works Scotland is publishing to share evidence, learning and ideas about public service reform. This paper relates in particular to the WWS **Improvement and Effectiveness** work stream.

**Professor Christopher Chapman** is Chair of Education and Public Policy and Practice, and Director of the Robert Owen Centre for Educational Change, University of Glasgow.

**Adele Rowe** is a Research Associate for the Robert Owen Centre for Educational Change, University of Glasgow.

*What Works Scotland is funded by the Economic and Social Research Council and the Scottish Government. [www.whatworksscotland.ac.uk](http://www.whatworksscotland.ac.uk)*

## Contents

1. Introduction.....	1
2. Scope and Context of Literature Review.....	2
3. Process Factors .....	7
4. Structural Factors.....	10
5. Conclusions.....	16
References.....	17
Annex.....	21

## 1. Introduction

In recent years there has been a major drive by both the Scottish and UK Governments to improve the quality of public services. Despite a number of Scottish Government initiatives post-devolution and significant growth in public spending, most indicators of social and economic inequalities have remained static or worsened and evidence shows that the social problems caused by inequalities place public services under increasing pressure (Christie, 2011). In 2010, the Scottish Government set up the Christie Commission with the remit of reviewing public service delivery in Scotland to examine challenges, obstacles and opportunities and recommend how public services might be reformed to improve outcomes for communities (Christie, 2011). The impetus for reform was triggered by the economic crisis that began in 2007 and resulted in budgetary pressures and lower revenues for Scotland. This was coupled with demographic and social changes that led to increased demand for services and a rise in public expectations (Donnelly, 2004):

“The daunting scale of that challenge is exacerbated by the expected drop in available funding for the next several years, the changes in Scottish demography, the resultant increase in demand, and an ever growing expectation of what public services should deliver.” (Christie, 2011 p2)

While governments have shown a clear commitment to the delivery of better public services, there is little discussion of what governments actually mean by public service improvement. The term is an elusive concept, framed alongside other terms closely associated with improvement such as effectiveness, efficiency, excellence, performance, reform and so on (Hodgson et al, 2007). Academic literature on defining public service improvement is also limited (Boyne, 2003a) yet there needs to be an explicit definition to establish a common understanding, along with robust empirical evidence on how public services can be improved (Hodgson et al, 2007).

Furthermore, individual public services have, to a greater or lesser degree, developed a shared language and understanding around the key definitions and concepts associated with improvement and effectiveness within their area of expertise; it remains a challenge to develop such shared perspective across the wider public services. We are reminded of conversations in the mid 2000's with directors of local authority services in England who had recently been reorganised from 'Directors of Education' to 'Directors of Children's Services'. Understandably those from a social services background tended to speak of reducing levels of neglect and abuse within the local authority and argue that if we created a more aware and caring community, outcomes, both social and educational would follow. Those from an educational background tended to argue that if educational attainment could be improved, this would lead to improved destinations and life chances of young people which in turn would have a positive impact on the wider community and... Clearly, these conversations

highlight a false dichotomy but they also illuminate the challenge of bringing together professionals (and others) from a range of backgrounds, each with a distinctive discipline-based identity, history and culture, career pathway, set of values and ultimately vocation to develop a shared inter-professional language and understanding, ultimately underpinned by a common set of values.

The dominant perspectives and biases that underpin and drive policy, strategy and practice can have major consequences. For example, the case of Haringey Council, involving Victoria Climbié's killing in 2000, followed by the death of "Baby P" in 2007 and subsequent use of special powers by the Labour Minister to remove the Director of Children's Services, Sharon Shoesmith<sup>1</sup> from her post are stark reminders of the high stakes involved for all concerned.

## 2. Scope and Context of Literature Review

Given the context outlined above, the purpose of this literature review is three fold. First, it will provide a starting point for framing the What Works Scotland position relating to the key concepts associated with Improvement and Effectiveness *across* the public services. Second it provides an opportunity to stimulate a debate about the key definitions and concepts *across* the public services and third, it will identify key areas for further investigation and review. With this in mind the review is organised around two key questions:

- How can we define public service improvement?
- What works in improving public services?

Therefore, the first part of this paper reviews the academic literature on public service improvement to provide a working definition of the term and considers the main theoretical models of improvement. In the second part of the paper, a sample of the empirical evidence on 'what works' in improving public services is presented to help judge which of many strategies is likely to lead to improvement in the delivery of public services.

### Question 1: How can we define public service improvement?

At the most fundamental level, improvement is defined as "an occasion when something gets better or when you make it better" (Cambridge Dictionaries Online). Defining improvement in the private sector is relatively straightforward as profit and shareholder value provide clear measures of success (Hodgson et al, 2007; Downe et al 2010). In the public sector, however, a definition is more complex as public services are concerned with social as well as fiscal outcomes (Hodgson et al, 2007) and have a far wider range of success criteria including "their effectiveness in meeting organisational goals; the efficiency with which they deploy resources; performance compared to other service providers; equality of

---

<sup>1</sup> Shoesmith was later awarded £680,000 for unfair dismissal see: <http://www.bbc.co.uk/news/uk-england-28454800>

access to services; robustness of their internal processes; and the transparency and accountability of decision-making processes” (Downe et al, 2010 p667).

Boyne (2003b) argues that definitions of performance and improvement in the public sector are ‘multi-dimensional, inherently political, dynamic and contested’ (Downe et al 2010) primarily because public service improvement is judged differently by different stakeholders, e.g. service users, politicians and tax payers. The logical consequence of this is that there can be no set criteria for determining whether or not a public service has improved. However, drawing on the literature on organisational performance in the public sector, Boyne (2003b) identifies a number of ‘tangible elements’ that all stakeholders are likely to view as desirable. These elements include: quantity of outputs (e.g. number of operations, hours of teaching) quality of outputs (e.g. reliability of service); efficiency (ratio of outputs to financial inputs); equity (fairness in distribution of service costs and benefits between different groups); outcomes (e.g. percentage of pupils passing exams); value for money (cost per unit of outcome) and consumer satisfaction. Boyne (2003b) reasons that an increase in any of the above (from the perception of stakeholders) can be viewed as evidence of improvement. He offers a working definition of public service improvement as: ‘a closer correspondence between perceptions of actual and desired standards of public services’ (Boyne, 2003a p223). Hodgson et al (2007) highlight that Boyne’s definition of improvement emphasises the importance of context:

“...improvement is contextually dependent on the politics of the day. Further it is spatially dependent on the country (and, arguably, the region/area and even organisation) under examination, historically dependent, hotly disputed and under-defined’.

While Boyne’s work provides a strong foundation for promoting a shared understanding of public service improvement, Boyne (2003a) himself highlights that his definition raises a number of important questions:

- “What criteria are used by different stakeholders to judge standards of public services?
- How much consensus is there on these criteria and their relative importance?
- How are actual and desired levels of performance on these criteria identified and by whom?
- To what extent is improvement an actual rise or fall in service standards?
- How closely are perceptions of improvement related to movements in objective indicators”?

### ***Models of improvement***

The paucity of literature on public service improvement is observed in a number of papers (e.g. Boyne 2003a; Hodgson et al 2007 and Downe et al 2010). Indeed, Downe et al (2010) notes that “the theories of improvement which influence organisational assessments,

performance frameworks and wider reform strategies are rarely made explicit by policy makers nor are they complete, comprehensive or entirely consistent systems of thought” (p667) yet they influence both policy and practice. Due to the shortage of academic literature on public service improvement, Boyne (2003a) draws on five conceptual models from the wider literature on organisational effectiveness to provide the theoretical basis for his working definition of improvement (see above). Before considering these models, it is important to note the two important differences that Boyne identifies between the concepts of effectiveness and improvement.

Firstly, while the effectiveness literature concentrates on the success or otherwise of individual organisations, the focus in public service improvement is on *services*, which are usually comprised of a set of organisations (e.g. all hospitals, schools or housing associations). Public service improvement may therefore require standards to be raised across a network of organisations for services to improve. Boyne (2003a) raises the important question about how public service improvement is measured: does a mean increase in the performance of a service indicate improvement even if standards in some organisations have remained the same or worsened? The issue of equity is also important: has improvement occurred if mean performance increases but the gap between best and worst organisations widens (Boyne, 2003a)?

Secondly, there is a difference in terms of the dimension of time. The literature on organisational effectiveness tends to focus on an organisation’s performance within a specific time frame, while public service improvement is concerned with shifts over time against a baseline position. In contrasting organisational effectiveness and public service improvement, Boyne stresses that improvement is a “dynamic phenomenon, inherently raises questions of equity and is usually concerned with the performance of multi-organisational networks rather than the achievements of single organisations” (Boyne, 2003a p213).

The five models discussed by Boyne (2003a) consider different criteria of organisational effectiveness and form the theoretical basis of his working definition of public service improvement (see above). The models are briefly summarised below.

### *The goal model*

The basis of the goal model is that organisations are established to achieve certain objectives that individuals, families, or communities cannot achieve alone (Boyne, 2003a). The success or failure of an organisation is then simply measured by the extent to which it achieves its aim and objectives, in the eyes of its stakeholders (Hodgson et al 2007). The simplicity of this model is attractive, however, there are practical problems in applying this model as public organisations often have a number of goals (some met, others not) and may not have well-defined and specific goals (Hodgson et al, 2007). There is also the problem of

attribution, i.e. disentangling cause and effect and controlling for other potential influences on outcomes (Boyne 2003a).

#### The systems-resource model

The systems-resource model is based on the idea of organisational survival: organisations that survive are successful by definition as market pressures ‘weed out’ the ineffective ones (Boyne 2003a). In this model, improvement is evident if an organisation is successful in gaining the resources (e.g. inputs such as services from suppliers) it needs to operate (Hodgson et al, 2007). The relevance of the model to the public sector, however, is questionable given that public organisations may be shielded from market forces or preserved (or abolished) for political reasons (Boyne, 2003a). The model raises the important question as to whether extra resources lead to improvement. In his review of the empirical literature, Boyne (2003a) finds little evidence of a link between spending and performance, but notes that most research is based on cross-sectional data that does not allow for the necessary analyses over time.

#### Internal processes model

The internal processes model uses organisational attributes as proxies for effectiveness, e.g. technical aspects such as information or budgeting systems, or human aspects such as human resource approaches that promote job satisfaction. Its relevance and appeal to public services is in using outputs as measures of improvement, rather than outcomes which are more difficult to define and measure. The model rests on the assumption that outputs are strongly related to organisational performance, however, Boyne (2003a) finds no evidence to support this view and argues that the “potential antecedents of service improvement are being confused with improvement itself” (Boyne, 2003a, p219).

#### The competing values model

This model assumes that all organisations face contradictory criteria in relation to their performance and attempts to combine elements of the previous three models (Boyne 2003a). Its relevance to public service improvement is in highlighting the contradictions between different views of effectiveness – e.g. the tension between ‘doing more with less’. It overlaps with the ‘balanced scorecard’ approach which emphasises the idea that no single indicator can be used to measure performance, but rather that multiple measures are necessary such as customer satisfaction, quality, service outputs and outcomes (Hodgson et al, 2007).

#### The multiple constituency model

The final model assumes that an organisation’s performance is judged by a variety of internal and external stakeholders, each of whom apply different criteria to assess



effectiveness and attach varying importance to the criteria (Boyne, 2003a). An organisation is effective to the extent that its stakeholders perceive it to be and its success is determined by its ability to mobilise and maintain political support (Boyne, 2003a). Central to the multiple constituency model is the idea that it is an organisation's image and legitimacy that matter most, rather than service delivery (Boyne, 2003a).

### ***From models of organisational effectiveness to a definition of public service improvement***

Boyne (2003a) offers a combination of the goal model, which stresses the *content* of improvement, and the multiple constituency model, which highlights the *process of improvement* - how and by whom improvement is defined, as the basis for defining public service improvement. Taken together, the models imply that a working definition of improvement must incorporate both the substance of an organisation's achievements and the political nature of judgments on its performance.

### **Question 2: What works in improving public services?**

Before delving into the empirical evidence on what works in public service improvement, it is important to reflect on the nature and quality of the available evidence base. Hodgson et al's (2007) review of academic literature on public service improvement in the UK post 1997 is based on 51 empirical studies, the majority of which focus on the NHS, education or local government, and include both quantitative and qualitative studies. In critiquing the quality of evidence, Hodgson et al (2007) describe a number of limitations which make it difficult to draw conclusions from the literature. The range of methods employed in the literature, the fact that much of the evidence is based on single case studies, and the different service contexts are problematic for making generalizations. For example, a finding that leadership is an important factor in schools cannot be generalized to other public sector organisations. There is the perennial problem of attribution whereby some studies assume that any upward shift in performance can be attributed to reform, without examining whether a causal link exists or whether the improvement might be explained by other factors. In a similar vein, Boyne (2003b) notes that the literature 'lacks rigorous causal reasoning and precise propositions'. Hodgson et al also highlight that many studies provide snapshots of improvement in a specific time frame but evidence is lacking as to whether changes are sustained over time. Lastly, (in line with the internal processes model above) studies often focus on improvements in *processes*, rather than outputs or outcomes, assuming that the former leads to the latter without being able to demonstrate this.

Overall, Hodgson et al (2007) conclude that the evidence on improvement in public services is limited, both in terms of quality and the range of sectors covered (most studies are based on schools, hospitals and local government) and note particular gaps in relation to the impact of strategic planning, strategic management and inspection.

With these limitations in mind, the remainder of the review seeks to summarise the empirical evidence on the recurring themes in the public service improvement literature and frame the improvement landscape. It is important to acknowledge that what follows is an overarching review of the evidence that seeks to identify those factors most commonly discussed in relation to public service improvement. Each of these factors warrants its own review to fully assess the evidence base and its impact on the delivery of public services.

### **3. Process Factors**

#### ***Improvement Science/Implementation Research***

The term 'Science of Improvement' or 'Improvement Science' is most commonly found in literature on health care and essentially refers to a scientific field concerned with which improvement strategies work. However, Perla et al (2003) note that the term is increasingly being used by a range of people and professions to mean different things. Langley et al were the first to use the phrase in the first edition of 'The Improvement Guide' (1996) which presented two central ideas. Firstly, that improvement rests upon developing, testing and implementing changes and measuring the impact of changes over time. Secondly, that subject matter experts are key to developing changes that are most likely to lead to improvement (Perla et al, 2003).

Improvement Science is then concerned with how knowledge in a specific field can be applied to different contexts. For example, can the means of improvements achieved in hospitals, schools, shops, insurance companies, communities and so on be successively transferred to other organizations? (Perla et al, 2003). If changes are informed by the knowledge and skills of subject matter experts and the results of changes are scientifically tested, proponents of Improvement Science believe the answer to this question is yes: "framing the change ideas suggested by subject matter experts using a scientific approach in a real world context is the essence of the science of improvement..." (Perla et al, 2003 p172).

#### ***Asset-based Approaches***

Asset based approaches originated in the sphere of health improvement and promotion but have rapidly spread to other fields, particularly community development, community work, social work, and criminal justice. Asset based approaches focus on identifying the strengths and capabilities of individuals or communities and harnessing these to promote positive change. The term 'asset' is all encompassing and refers variously to skills, resources, knowledge, social capital, physical and environmental resources. The approach stands in direct contrast to a needs based approach or deficit model which concentrate on weaknesses and 'fixing' problems and are associated with 'doing to' rather than 'working with'. The rationale for focusing on strengths is to empower and enable communities or individuals to identify what works well and to use this as the basis for galvanising action.

“A growing body of evidence shows that when practitioners begin with a focus on what communities have (their assets) as opposed to what they don’t have (their needs) a community’s efficacy in addressing its own needs increases, as does its capacity to lever in external support. It provides healthy community practitioners with a fresh perspective on building bridges with socially excluded people and marginalised groups.”

(Foot and Hopkins, 2010, p6)

Empowering communities to take control of their own resources (Foot and Hopkins 2010) and lead their own development are central to the approach as proponents argue that this will create more equitable and sustainable economic and social development at community level (Mathie, 2006). Critics of the approach, however, argue that it is inadequately defined (Rapp et al., 2005) and that it does not differ significantly from other similar approaches in social work (Curtis McMillen, Morris and Sherraden, 2004). Despite this, the term has gained currency and the approach has become increasingly popular, particularly in Scotland where it is viewed as a central plank of Scotland’s approach to reforming public services.

Evaluations of asset based approaches tend to be qualitative and focus on thick descriptions and case studies that are well suited to the context specific nature of the approach. There is a lack of empirical evidence on effectiveness, which in part seems to stem from a view that empiricism does not sit well with asset based approaches.

“In practice this results in evaluation reports focusing on ‘traditional indicators’ such as ‘number of participants (reach) and individual changes in knowledge or behavior’, yet, unable to evidence changes using such measurements, resort to persuasive anecdotal accounts due to this methodological impasse” (Hills et al., 2010, p96).

Evidence from case studies suggests that asset based approaches can have a positive impact on communities (McGuinness, 2015). An asset based regeneration project in two neglected housing estates in Falmouth, Cornwall mobilised local residents to form a tenants committee that delivered a newsletter and visited all households in the estates to inform residents of their plans for the estate. A series of community meetings were held and from these it was ascertained that the main problems affecting the estates were crime, poor housing and unemployment, coupled with a failure on the part of statutory agencies to take action. The project resulted in the creation of a resident led Beacon Community Regeneration Partnership which won grants to insulate houses and install central heating. Other indicators of success included reduction in crime rates, a fall in the number of children on the Child Protection Register, and improved exam results (O’Leary et al., 2011).

### ***Collaborative practices/Partnership working***

Collaborative practices and partnership working feature strongly in the Christie Commission’s (2011) vision of public service reform as a ‘radical, new collaborative culture’.

Indeed, one of the Commission's key recommendations is that public service organisations in Scotland should work more collaboratively to 'extend and deepen' partnership approaches and go beyond the current model of community planning partnership (Christie, 2011). Collaborative working is viewed as a means of bringing together the patchwork of organisations from the public, private and third sectors involved in delivering public services to promote joint working, minimize duplication and deliver more integrated services. Central to the principles of collaborative practices is working closely with individuals and communities to understand their needs and strengths and involving them in the design of services in a 'bottom-up' approach.

Collaborative working can be seen as a response to the fragmentary effects of the 1980s New Public Management approaches (see below) in that it brings different parts of the system back together (Lamie and Ball, 2010). It is viewed positively by many involved in local government in Scotland (Lamie and Ball, 2010) and it is difficult to argue against what seems to make good intuitive sense. Collaboration, however, carries with it costs in terms of time and resources, which means that evaluation is vital to establish whether or not it has a positive impact on public services (Lamie and Ball, 2010). Despite the Scottish Government's drive towards greater collaborative working, there is little empirical evidence to judge whether or not collaborative practices actually lead to public service improvement.

One such study concerns an evaluation of partnership working within the context of community planning and was carried out by a local council in the east of Scotland. The study used qualitative interviews with the partners, focus groups with residents' associations, participant observations and a 'partnership assessment tool' to gauge success. From the partners' perspectives, the partnership was working reasonably well but could do better. In contrast, the local residents group responded negatively to a wide range of aspects and generally perceived that service delivery had not improved (Lamie and Ball, 2010).

Collaborative practices in the field of education are now commonplace in the UK and school-based networks, collaborations and federations are viewed as an important means of improving a school's educational offer. Hadfield and Chapman (2009) define school networks as 'educational endeavours entered into voluntarily by staff and pupils' (p2) and emphasise the voluntary nature of collaborations as active participation is key to their success. Before considering the evidence base on the impact of school networks on pupil outcomes, Hadfield and Chapman (2009) note the particular difficulties in attributing causality. Networks are based on multiple connections and a variety of interactions that make it especially difficult to disentangle the impact of one network activity in amongst a multitude of others. Given this difficulty, much of the literature focuses on *how* networks impact on schools in relation to their effects on leaders, staff, pupils via a range of intermediate or 'proxy' outcomes such as changes to teachers' attitudes and practice (Hadfield and Chapman, 2009). In the UK there have been a number of school-based

networking initiatives that have been funded by central government and focused on either poorly performing schools or highly successful schools (Hadfield and Chapman, 2009). A review of 17 different UK networks focused on inner cities or challenging circumstances found that “well-led and appropriately structured collaboration between schools facing complex and challenging circumstances helped their leaders to balance short term pressures to improve pupil attainment with long term desires to improve the educational experiences of their pupils and the engagement of their communities” (Hadfield and Jopling, 2007, p3). The findings from Hadfield and Jopling’s 2007 review suggested that collaboration had helped schools take a longer-term approach to raising attainment rather than opting for quick wins.

#### **4. Structural Factors**

##### ***Improvement through leadership and management***

The importance of leadership and management is a recurring theme in the improvement literature. The Audit Commission (2002) identified ‘poor leadership’ as a major cause of service failure in local government. Examples of changing leadership in a bid to improve public services include the appointment of ‘superheads’ in failing schools, and ‘interim managers’ in local government (Boyne, 2004). Hodgson et al’s (2007) literature review found 11 studies that reported a positive impact for leadership and management change on public sector improvement. However, all but one of these studies was based on a qualitative approach and relatively small scale. A common finding was that new leaders brought about improvements in processes, such as a change in culture, greater collegiality or new ways of working and helped raise expectations and increase staff motivation. For example, Hadyn’s (2001) case study of a comprehensive school in England posits a link between a new head teacher, better exam results and improvements in the Ofsted report (Hodgson et al, 2007).

The relationship between leadership and effective schooling is reported in a number of studies and is an important strand of inquiry in the literature on educational effectiveness and improvement (Mongon and Chapman, 2012). Day et al’s (2009) study for the Department for Children Schools and Families looks at effective school leadership and pupil outcomes. Focusing on a sample of schools that had raised pupil attainment over three consecutive years and under the leadership of the same head teacher, the mixed methods study found that “heads in more effective schools are successful in improving pupil outcomes through who they are – their values, virtues, dispositions, attributes and competences – the strategies they use, and the specific combination and timely implementation and management of these strategies in response to the unique contexts in which they work” (Day et al, 2009, p1). The study found statistically significant and ‘qualitatively robust’ associations between pupil outcomes and heads’ educational values, qualities and strategies in improving school conditions. Head teachers in the study were

found to possess a common set of educational values (based around promoting individual and social well-being and raising attainment for all pupils), personal qualities, dispositions and competencies. These, combined with strategies the heads developed that responded to individuals' and their school's needs and took account of their school's unique context, produced cultural change and changes in teaching practices that were influential in improving student outcomes (Day et al, 2009).

### ***Improvement through New Public Management practices***

The term New Public Management (NPM) was formally conceptualized by Hood (1991) and refers to a range of government policies introduced in the 1980s that aimed to modernise and increase efficiency in the public sector by following market oriented management principles. It is used to describe a way of 'reorganizing public sector bodies to bring their management, reporting, and accounting approaches closer to (a particular perception of) business methods' (Dunleavy and Hood, 2009 p9). Bezes et al (2012) outline five organizational principles associated with NPM: firstly, separating strategy, leadership and control from implementation; secondly, fragmenting vertical bureaucracies by setting up independent administrative agencies; thirdly, systematically resorting to market mechanisms; fourthly, transforming bureaucratic hierarchies by reinforcing the responsibilities of officials in charge of implementing state action; and lastly, introducing result management based on achieving objectives and measuring performance.

Market competition is believed to promote efficiency, innovation, consumer choice, and responsiveness and traditional economic theory suggests a positive relationship between competition and service performance. 'Outsourcing' services to market suppliers, greater engagement with the voluntary sector, the blurring of boundaries between public and private sectors and increasing the 'devolution of responsibility to individuals, families and households to provide for their own well-being' are claimed to drive improvements in efficiency and quality (Clarke, 2008). However, the impact of markets in public services is hotly debated. Clarke (2008) argues that such practices can create a fragmented system that makes co-ordination more difficult. Furthermore, competition is criticized for its adverse effect on equity in that disadvantaged groups are likely to lose out as they are more expensive to treat but most in need of public services (Le Grand and Bartlett 1993).

Performance indicators, frameworks and monitoring are core features of NPM and have been widely used in the UK as a means of driving improvement across public services including schools, hospitals, social care providers, local government, police and the probation service (Downe et al, 2010; Hodgson et al, 2007). The theory behind performance measures and targets is that they help organizations to focus on outcomes rather than processes, they provide direction, and they motivate people to improve performance (Boyne and Chen, 2007). Critics, however, argue that they hinder performance, encourage 'gaming', and result in the neglect of other areas (Boyne and Chen, 2007).

Performance frameworks are defined as “interconnecting systems of reporting arrangements, resource flows, performance measures and organisational assessments and interventions that are designed to steer the performance of public service organisations” (Downe et al, 2010 p666). For example the Public Service Agreements (PSAs) introduced under New Labour focused on improving effectiveness and efficiency through setting aims, objectives and targets. Local PSAs provided additional funding for local authorities that were successful in meeting a set of performance targets that were negotiated with central government (Hodgson et al, 2007). While the empirical evidence suggests that setting clear goals is positively related to organisational performance, there is a lack of evidence on the specific question as to whether or not targets lead to better performance in public services leading Boyne and Chen (2007) to describe local PSAs as an ‘unprecedented experiment in payment by results’ (p456).

Boyne and Chen’s 2007 study seeks to address this gap by analyzing the impact of local PSA targets on education performance (exam results) in 147 English local authorities between 1998 and 2003 to assess whether performance is better or worse in organizations that set quantifiable targets. They find that local authorities with a target performed better than their counterparts without a target, controlling for baseline performance. While this suggests that target setting can lead to better outcomes, Boyne and Chen rightfully point out the limitations of their study which deals with only one target regime in a sub-set of English LAs and focuses specifically on the impact on exam results. They raise wider questions about the impact of targets on equity and value for money in public services.

Downe et al (2010) compare and contrast the development of performance frameworks in Scotland, England and Wales to understand why and how the different parts of the UK have adopted varied approaches to monitoring local government performance. Comprehensive Performance Assessment (CPA) in England, introduced in the 2001 White Paper, was based on bringing together information held by government departments, auditors and inspectors on each council into a single framework. CPAs were explicitly designed to encourage competition and comparison between local authorities, based on the premise that poorly performing LAs were unaware of their weaknesses and required external pressure to address them. Similarly, Scotland’s Best Value Audits were designed to provide an impetus for improvement and were viewed as ‘report cards’ on local authority Chief Executives. Although they also provided external challenge, they sought a more consensual approach between councils and auditors. At the other end of the spectrum, Wales rejected performance comparison on the grounds that league tables encouraged ‘gaming’ and improvement had to come from within. Instead, auditors worked with Welsh councils to tailor assessments to each authority, making it difficult to draw comparisons between areas. Based on interviews with senior policy makers, Downe et al (2010) find that all three approaches were credited with improving local councils’ capacity, particularly for the worst performers. CPAs were seen as having had the greatest impact, but some felt that their influence lessened over time as some LAs learnt how to ‘play the game’. Downe et al (2010)

argue that approaches taken to performance assessment do not only measure performance but also contribute to definitions of performance, e.g. CPA scores in England gained widespread acceptance of how English local authorities' performance should be judged.

### ***Improvement through inspection***

Closely linked to performance frameworks, targeting and monitoring is the inspection of public services. Over the last two decades inspection has been used increasingly to hold to account schools, hospitals, social care, local government, the police and other public service providers (Martin and Davis, 2008). Advocates of inspection view it as an essential tool for accountability and a driving force for improvement. Downe (2008) describes a range of theoretical benefits that the 'threat' of inspection may have: service managers may take greater care in their work; it can help to ensure minimum standards; it can protect service users by identifying failing services; and diagnoses of problems may mean that inspectors can help provide a remedy and ensure mistakes are not repeated elsewhere. Its critics, however, argue that inspection places a 'huge bureaucratic burden' on services and results in resources being shifted away from frontline services (Martin and Davis, 2008).

Despite the increasing importance of inspection, Martin and Davis (2008) note that there is relatively little academic literature on its operation and impacts. Assessing the impact of inspection is hindered by the problems of attribution referred to above in that it is extremely difficult to isolate its effects from other factors that contribute to improvement. However, as Downe (2008) points out, the costs of inspection have increased along with its use, raising the important question as to whether its benefits outweigh its costs.

The evidence on the impact of inspection in local government is mixed. The Audit Commission pointed to an improvement in Comprehensive Performance Assessment scores in England as evidence that the inspection regime led to improvement (Downe, 2008). However, others highlight the multi-faceted nature of performance and argue that aggregate measures like CPA fail to capture these complexities and ignore service users' experience (Downe, 2008). Research in Wales found that inspection had helped poorly performing authorities to improve (Lewis et al. 2007) and the Best Value Audits in Scotland were found to have a positive impact on most councils, yet there was no evidence of impact on service users (Downe et al., 2008).

Martin (2008) usefully clarifies the role of inspection in educational institutions as focusing on the "extent to which the institution is effective in ensuring high standards of pupil performance and achievement – as well as the capacity of the institution to improve the institutional standards of teaching, management, leadership, governance and teaching" (pp57-58). Similar to local government, evidence on the impact of inspections on schools is mixed and inconclusive. Ehren et al's (2014) recent review of the international evidence on school inspections finds that Ofsted has a 'powerful influence' on schools but whether this



influence is positive or negative depends on the type of school and the quality of inspection. The literature also suggests that the impact of inspection depends on the quality of feedback; whether or not inspection reports, test results and league tables are published; whether or not there is parental choice; the quality of school leadership and a school's capacity to improve; and the strength of sanctions and support (Ehren et al, 2014). Concerns with Ofsted inspections include variability in the quality of inspection teams, questions over the reliability of judgments and detrimental consequences of inspection such as teaching to inspection standards and manipulating inspection data (Ehren et al, 2014). A survey of teachers in Chapman's 2001 study on the impact of Ofsted on inspections on classroom change found that 51 percent of teachers felt that inspectors did not get a realistic picture of their teaching. Several studies highlight a tension between the benefits of inspection in providing objective feedback and informing strategies for improvement, and the need for a supportive process that helps to effect change (Martin, 2008).

Shaw et al. (2003) find that Ofsted inspection had no positive effect on exam performance and may have had a detrimental effect. Similarly, a study in 2004 by Ofsted and the Institute of Education found no effect on GCSE results noting that "inspection is neither a catalyst for instant improvement in GCSE results nor a significant inhibitor" (p31). However, Allen and Burgess' (2012) analyses finds a link between failing an inspection and improved student performance over the two to three years following inspection. Ofsted's 2004 report suggests that inspection contributed to improvement in all sectors (schools, early years provision, children's services), but most notably in schools and initial teacher training. The report acknowledges that alongside inspection, other developments in the curriculum, assessment, leadership and management of institutions and national strategies and standards for teaching had contributed to improvement. In particular, inspection was credited with helping to embed these initiatives by providing 'evaluation, leverage and accountability' (Mathew and Sammons, 2004).

A more recent evaluation of Ofsted (Ofsted 2007a) acknowledged that the relationship between inspection and improvement is not straightforward and it cannot be assumed that the former leads to the latter. The report stresses that it is the leaders, staff, children and young people – the service providers and service users – who actually bring about improvement with inspection acting as a 'catalyst' by helping institutions to gauge their own performance and direct improvement. In support of this, a survey conducted by the National Foundation for Educational Research found that nearly two thirds of respondents and just over half of case study interviewees felt that inspection had contributed to school improvement by confirming and clarifying areas for improvement (McCrone et al 2007). Chapman's 2001 study found that 55 percent of teachers viewed the feedback they received on their teaching as part of an inspection as useful, yet few (22%) felt they would actually change their practice as a result of the inspection. Ofsted's 2007 report also finds that inspection had an impact on the system as a whole by contributing to national policy, spending decisions and policy priorities.

### ***Improvement through resourcing***

Perhaps the most straightforward proposition in considering what might trigger public service improvement is that more resources will lead to better results. However, Boyne (2003b) finds that the majority of empirical evidence reveals no significant relationship between either financial resources or real resources and service performance. Boyne (2003b) notes that empirical studies looking at the impact of resources typically include measures of financial resources and real resources (staff, equipment etc.) in a *direct* effects model and argues that an *indirect* effects model is more plausible whereby financial resources allow real resources to be purchased and the latter influence performance.

Hodgson et al's (2007) review includes four studies that consider the effect of resources on service improvement: two examine the impact of financial resources in the NHS while the other two look at the impact of real resources in schools. Tsai and Mar Molinero's (2002) study found that efficiency in hospitals was improved where hospitals were able to allocate money to what they do best and Fiscbacher and Francis's (1998) qualitative account of GP fund holding found improvement in 'direct, open or rapid access services'. Turning to schools, Bradley et al's (2001) study found small positive correlations between resources (teachers, books, equipment) and both exam performance and increased school attendance. Mancebon and Mar Molinero (1998) found a link between teacher/pupil ratio and the percentage of successful English and Science test results for primary school age children.

More recently Nicoletti and Rabe's 2014 study examines the effect of school spending on pupils' test scores in English, Maths and Science in GCSE examinations and finds that most pupils benefit from more money spent on learning resources such as books, software and computers, but pupils with English as an additional language and some FSM pupils do not seem to benefit as much. Overall, the empirical evidence on the impact of resources on public service improvement is mixed and Hodgson et al note that the literature tells us little about *how* resources lead to improvement.

### ***Improvement through (re-)organisation***

The importance of the structure and size of an organization for public service improvement is another factor considered in the literature. Hodgson et al's (2007) review includes a number of studies of interest. Bradley et al's (2001) account of 2657 secondary schools in England indicates improvements in exam performance and attendance rates as the size of schools increase. McCallion et al's (1999) examination of 23 hospitals in Northern Ireland between 1989 and 1992 found that larger hospitals displayed higher levels of 'cost efficiency, higher allocative efficiency and higher technical efficiency than their smaller counterparts' (1999, p.27). And Cereste et al's (2003) examination of NHS mergers found the PCTs which had merged were seen to be more effective in reconfiguring services with

improvements in patient care. However, a study examining the effects of size on efficiency within 44 police forces in England and Wales found that the smaller the force, the greater the efficiency. This finding was largely attributed to the fact that larger orgs have extra layers of bureaucracy and management that hinder efficiency (Drake and Simper, 2000).

Boyne (2003b) notes that reorganisation is a common feature when redesigning public service provision that is often more about symbolic change than tangible benefits. A recent example may be when the incoming UK coalition government came to power in 2010 and made changes to a number of Whitehall Departments including rapidly renaming the Department for Children Schools and Families the Department for Education and issuing a new logo to signal the ethos and policy priorities of the incoming Government.

## **5. Conclusions**

In seeking a better understanding of public service improvement this review has provided a definition of the term, described the main theoretical models and framed a number of processes and structures that are identified in the literature as contributing to the improvement of public services. The range of topics included in the second part of the review is testament to the breadth of factors that warrant study when investigating 'what works' in relation to public service improvement. The range of issues, however, stands in contrast to limitations with respect to the quality and extent of the evidence in some areas. A lack of robust studies that demonstrate causality, and/or the difficulty in attributing impact and establishing causality is highlighted in many studies. The summaries of evidence included in the second part of the review provide only a brief insight into the literature on each topic. However, it is hoped that this goes some way in framing the public service improvement landscape and helping to identify a number of areas that warrant further investigation and more in depth work by What Works Scotland. In addition to the What Works Reviews on 'Spread and Sustainability' and 'Asset Based Approaches' the next phase of activity will involve undertaking detailed reviews on specific aspects raised within this paper. The first of which will be a review of evidence on Improvement Science as a lever for improvement within the public services.

## References

- Allen, R. and Burgess, S. (2012), *How should we treat under performing schools? A regression discontinuity analysis of school inspections in England*, Centre for Market and Public Organisation Working Paper No. 12/287.
- Audit Commission (2002). *A Force for Change: Central Government Intervention in Failing Local Government Services* London: HMSO
- Bezes, P., Demaziere, D., Le Bianic, T., Paradeise, C., Normand, R., Benamouzig, D., Pierru, F. & Evetts, J. (2012). New public management and professionals in the public sector. What new patterns beyond opposition? *Sociologie Du Travail*, 54 E1-E52.
- Boyne, G. A. (2003a). What is public service improvement? *Public Administration*, 81(2) 211-227.
- Boyne, G. A. (2003b). Sources of public service improvements: A critical review and research agenda. *Journal of Public Administration Research and Theory*, 13(3) 367-394.
- Boyne, G. A. (2004). 'A '3Rs' strategy for public service turnaround: Retrenchment, repositioning and reorganization.' *Public Money & Management*, 24(2) 97-103.
- Boyne, G., Martin, S., & Walker, R. (2004). Explicit reforms, implicit theories and public service improvement. *Public Management Review*, 6(2) 189-210.
- Boyne, G.A. & Chen, A.A. 2007; 2006, Performance Targets and Public Service Improvement, *Journal of Public Administration Research and Theory*, 17(3) 455-477.
- Bradley, S., G. Johnes & Millington J. (2001). The Effect of Competition on the Efficiency of Secondary Schools in England. *European Journal of Operational Research*, 135 545–68.
- Cereste, M., N. Doherty and C. Travers (2003). An Investigation into the Level and Impact of Merger Activity amongst Hospitals in the UK's National Health Service. *Journal of Health Organization and Management*, 17(1) 6–24.
- Chapman, C. (2001) Changing classrooms through inspection. *School Leadership and Management*, 21(1) 59-73.
- Christie (2011). *Report on the Future Delivery of Public Services by the Commission* chaired by Dr Campbell Christie. <http://www.scotland.gov.uk/Publications/2011/06/27154527/0>
- Clarke, J. (2008) Performance Paradoxes. The Politics of Evaluation in Public Services. In H. Davis and S. Martin (eds) *Public Services Inspection in the UK, Research Highlights 50*. London: Jessica Kingsley Publishers.

Curtis McMillen, J., Morris, L. and Sherraden, M. (2004) Ending Social Work's Grudge Match: Problems versus Strengths, *Families in Society: The Journal of Contemporary Social Services* 85(3) 317-325

Day, C., Sammons, P., Hopkins, D., Harris, A., Leithwood, K., Gu, Q., Brown, E., Ahtaridou, E., Kington, A. (2009). *The impact of school leadership on pupil outcomes: final report*. Nottingham: DCSF. DCSF Research Report 108  
[http://www.almaharris.co.uk/files/the\\_impact\\_of\\_school\\_leadership.pdf](http://www.almaharris.co.uk/files/the_impact_of_school_leadership.pdf)

Donnelly, M. (2004) Cusp not crisis: Changing to Deliver Scottish Government. *International Journal of Productivity and Performance Management*, 53 (6) 548-557.

Downe, J., Grace, C., Martin, S. J., & Nutley, S. M. (2010) Theories of public service improvement. A comparative analysis of local performance assessment frameworks. *Public Management Review* 12 (5) 663-678

Downe, J. (2008) Inspection of Local Government Services. In S. Martin and H. Davis (eds) *Public Services Inspection in the UK, Research Highlights 50*. London: Jessica Kingsley Publishers.

Downe, J., Grace, C., Martin, S. J., and Nutley, S. M (2008). Best Value Audits in Scotland: Winning without scoring? *Public Money and Management* 27,2.

Drake, L. and R. Simper (2003). The Measurement of English and Welsh Police Force Efficiency: A Comparison of Distance Function Models. *European Journal of Operational Research*, 147 165–86.

Dunleavy, P. and Hood, C. (1994). From old public administration to new public management. *Public Money & Management*, 14(3) 9-16

Ehren, M, Perryman, J & Spours, K, (2015). *Accountability and school inspections*. Essay for 2015 UK General Election, 2014.  
[http://www.ioe.ac.uk/about/documents/About\\_Policies/Ehren\\_et\\_al\\_-\\_FINAL.pdf](http://www.ioe.ac.uk/about/documents/About_Policies/Ehren_et_al_-_FINAL.pdf)

Ferlie, E, Ashburner, L, Fitzgerald, L, & Pettigrew, A. (1996). *New Public Management in Action*. Oxford: Oxford University Press.

Fiscbacher, M. and Francis, A. (1998). Purchaser Provider Relationships and Innovation: A Case Study of GP Purchasing in Glasgow. *Financial Accountability and Management*, 14(4) 281–98.

Foot, J., & Hopkins, T. (2010). *A glass half-full: how an asset approach can improve community health and well-being*. Great Britain Improvement and Development Agency.

Hadfield, M. and Chapman, C. (2009). *Leading School-based Networks*. Oxon: Routledge.

- Hadfield, M. and Jopling, M. (2007). *The Potential of Collaboratives to Support Schools in Complex and Challenging Circumstances*. Nottingham: NCSL.
- Haydn, T. (2001). From a Very Peculiar Department to a Very Successful School: Transference Issues, School Leadership and Management, 21(4) 415–39.
- Hodgson, L, Farrell, C. M. & Connolly, M. (2007) Improving UK Public Services: A Review of the Evidence. *Public Administration*, 85 355–382.
- Hood, C. (1991). A Public Management for All Seasons. *Public Administration*, 69 3-19.
- Lamie, J. & Ball, R. 2010, Evaluation of Partnership Working within a Community Planning Context, *Local Government Studies*, 36 (1) 109-127.
- Langley, G. J, Moen, R, Nolan, K. M, Nolan, T. W, Norman C. L, Provost, L. P, (2009). *The Improvement Guide: A Practical Approach to Enhancing Organizational Performance*. New York: John Wiley & Sons.
- Le Grand, J. and Bartlett, W. (eds) (1993). *Quasi-markets and social policy*. Basingstoke: Macmillan Press.
- Lewis, M., Phillips, W., and Percy, N. (2007). *Codifying Local Government Support and Intervention*. Cardiff: National Assembly for Wales.
- Martin, J. (2008) Inspection of Education and Skills: From Improvement to Accountability. In S. Martin and H. Davis (eds) *Public Services Inspection in the UK, Research Highlights 50*. London: Jessica Kingsley Publishers.
- Martin, S., and Davis, H. (2008) The Rise of Public Services Inspection. In S. Martin and H. Davis (eds) *Public Services Inspection in the UK, Research Highlights 50*. London: Jessica Kingsley Publishers.
- Mathews, P. and Sammons, P. (2004). *Improvement through inspection: An evaluation of the impact of Ofsted's work*. HMI 2244. London: Ofsted.
- McCallion, G.J. McKillop, C. Glass and C. Kerr (1999). Rationalizing Northern Ireland Hospital Services Towards Larger Providers: Best-practice Efficiency Studies and Current Policy, *Public Money & Management*. April–June, 27–32.
- McCrone, T., Rudd, P., Blenkinsopp, S., Wade, P., Rutt, S. and Yshanew, T. (2007). *Evaluation of the Impact of Section 5 Inspections*. Slough: NFER.
- McGuinness (2015) ABCD: A literature Review (unpublished manuscript)

Mongon, D. and Chapman, C. (2012). *High-Leverage Leadership: Improving Outcomes in Educational Settings*. Oxon: Routledge.

Nicoletti, C. & Rabe, B. (2015) *Spending it wisely: How can schools use their resources to help poorer pupils?* Nuffield Foundation:

<https://www.iser.essex.ac.uk/files/news/2014/spending-it-wisely.pdf>

Ofsted (2007a) *Review of the Impact of Inspection*. London: Ofsted.

O'Leary T, Burkett I, Braithwaite K. *Appreciating Assets*. Carnegie UK Trust and International Association for Community Development. Carnegie UK Trust: 2011

Perla, R. J, Provost L.P, Parry, G. J, (2013) Seven propositions of the Science of Improvement: Exploring Foundations. *Quality Management in Health Care*, 22(3) 170.

Rapp, C.A., Saleebey, D. Sullivan, W.P (2005) The future of strengths-based social work. *Advances in Social Work*, 6(1) 79-90

Shaw, I., Newton, D.P., Aitkin, M., and Darnell, R. (2003). Do Ofsted inspections of secondary schools make a difference to GCSE results? *British Educational Research Journal* 29(1) 63-75.

Tsai, P.F. & Molinero, C. M, (2002). A Variable Returns to Scale Data Envelop Analysis Model for the Joint Determination of Efficiencies with an Example of the UK health Service. *European Journal of Operational Research*, 141 21–38

## Annex 1 – Methods

The review seeks to address the following questions:

- How is improvement and effectiveness in relation to public service delivery defined in the literature and what theories and models have been developed to explain improvement?
- ‘What works’ in terms of improving public service delivery in the UK?

To ensure the review is transparent and explicit about what it does and does not cover, explicit inclusion criteria and search strings were developed and used consistently in searching and sifting the literature.

Studies that met the following criteria were considered for inclusion in the review:

- peer reviewed;
- address improvement or effectiveness in public service delivery – from a theoretical or empirical standpoint (all public services included – e.g. health, social work, police, education, community services, prison and probation services);
- UK based studies (though reviews include international studies);
- English language studies;
- studies published since 2000.

The search strings: “public services” AND (improv\* or effectiv\*) and “public services” AND (improv\* or effectiv\*) AND Scotland were run on Social Services Abstracts, World Wide Political Science Abstracts and Business Source Premier and in total resulted in over 800 results. In addition, the search strings “improvement science” AND Scotland and “improvement science” were run on the same three databases as well as Applied Social Sciences Index and Abstracts (ASSIA), ERIC, and the International Bibliography of the Social Sciences (IBSS).

Due to time constraints, it was not possible to sift through the entire search results. Rather, studies that met the inclusion criteria and appeared most relevant to the review questions were ‘cherry picked’. It is important to acknowledge this limitation and the potential bias it poses.