
2 From the improvement end of the telescope: benchmarking and accountability in UK local government

Clive Grace¹
Cardiff Business School

2.1 Introduction

In the summer of 2010 the UK's newly elected coalition government announced the abolition of the principal benchmarking and performance management regime for local government in England, the Comprehensive Area Assessment, and its intention to abolish the principal authors and stewards of that regime, the Audit Commission (DCLG 2011) as well. The government also announced new requirements for public services to publish more information so that an 'armchair auditors' would be sufficiently equipped to hold those services to account directly, and without the intervening agency of bodies such as the Audit Commission.

These policies were introduced in the context of the wider programme of the coalition government with its emphasis on 'localism' and on the 'Big Society', and a comprehensive assault on the many intermediary and 'arms length bodies' such as the Audit Commission which were seen as fogging the relationship between government and citizenry. They were also a reaction to a decade or more of what was seen as top down performance management, inhibiting the exercise of professional discretion at the front line and creating a bureaucratic morass in which

¹ Clive Grace is currently an Honorary Research Fellow at Cardiff Business School and, previously, was a UK local authority CEO, and Director General of the Audit Commission in Wales. He was also Deputy Chair of the Beacons Scheme from 2003 to 2007. In consequence he had direct involvement and responsibility in respect of a number of the benchmarking and related programmes discussed in this piece and, accordingly, more than the usual caveats apply to the judgements expressed therein.

performance targets distorted behaviour and generated gaming as much as they generated genuine service improvement.

Yet the simultaneous celebration of the new ‘armchair auditors’ and sources of performance benchmarking information on local government services on which they might draw, such as that compiled by the Association of Public Service Excellence, gave more than a hint that the benchmarking era was changing rather than concluding. Indeed, it underlined the ubiquitous role and use of benchmarking in UK local government and other public services.

There are many varieties. First, there is a wide range of **service-based cost and technical comparisons** conducted as benchmarking ‘clubs’ of one kind or another. These include those of the Association of Public Service Excellence (APSE), a not-for-profit voluntary body established with service comparisons of ‘blue collar’ local government services as a core aim; the Chartered Institute for Public Finance and Accountancy (CIPFA), a major professional accountancy body for, *inter alia*, local government finance staff; and the Wales Audit Office (WAO), the statutory public audit body for Wales. Also in this area are the ‘communities of practice’ established across a range of different services by the Improvement and Development Agency (IDIA), an agency of the Local Government Association (LGA), which has now been absorbed within the LGA.

Secondly, there have been a series of centrally determined **performance indicator sets** with results often published in the form of league tables. Then there have been **performance regimes** for local authorities, looking at the whole organisation and testing them against pre-set frameworks, including the Comprehensive Performance Assessment (England), the Wales Programme for Improvement, and Best Value Audits (Scotland). These led in England to a yet wider programme of Comprehensive Area Assessments, which brought together data on a much wider group of local services. Alongside these performance regimes has been a programme of ‘voluntary’ assessments using **external peer review** methods against a framework underpinned by the European Framework for Quality Management. There have also been major **excellence benchmarking schemes**, which test projects and services against a pre-designed benchmark to identify best and excellent practice, and most notably the central government run Beacon Council Scheme in England.

These were (and still are, in some cases) all major programmes of work, and it all adds up to an awful lot of benchmarking, and to considerable direct and indirect cost. Perhaps unsurprisingly, therefore, it all figured strongly in the unfolding dynamic between local and central government amongst the constant striving for

improved local government services and performance that started under Thatcher but which gathered even greater pace under the Blair–Brown Labour governments.

There are good (structural) reasons why benchmarking in UK local government has been so widespread. Accordingly, **Section Two** briefly sketches out key features of UK local government along with an indication of the position in respect of the devolved governments of Wales and Scotland, and then describes the range and scope of UK local government benchmarking and why it matters. **Section Three** shows how the popularity and use of the various benchmarking methods has ebbed and flowed across the period of 1998 to 2010 as an expression of three distinct eras of local–central relationships, all of which were (and are) concerned, broadly, to achieve improved services and stronger accountability. Each of these eras has associated with it not only particular instruments to secure that improvement but also a ‘theory’ of improvement — sometimes explicit and sometimes implicit, and more or less comprehensive and credible — as to how those instruments might help achieve the desired outcome.

In **Section Four** we look in more detail at a selection of the benchmarking methods that have been deployed in and across these different eras — including such assessment as is available of the efficacy of the various rewards and punishments that they have employed to try and secure better services. In conclusion, **Section Five** draws out the underlying fundamental relationship that connects benchmarking to service improvement, and suggests a rudimentary framework through which policy makers and practitioners should approach the use of benchmarking methods. It also expresses some tentative views on the application of that framework to securing improvement in public services in an age of austerity, that being firmly the context of public services and their improvement in the UK for the foreseeable future.

In doing all this, the aim of the chapter is threefold. First it provides its own ‘benchmark’ of comparison as between the decidedly unitary governmental system of the UK and the divided governance characteristic of federal systems (see Fenna, this volume). This is not about ‘learning lessons from the UK’, but simply about providing a broader landscape of contrast and difference to help generate insight and understanding. Secondly, the creation of devolved governments covering the 13 per cent of the UK population in Scotland and Wales adds further to the comparative mix. It does not of course make the UK ‘federal’ in any fundamental sense. As yet, neither the devolved governments nor indeed any of the local government units in any part of the UK constitute truly different ‘orders of government’, each with their own separate source of constitutional validity and powers. But UK devolution has created lines of policy tension and divergence that can augment the value of UK-related comparison. Thirdly, there has been so much

performance benchmarking of UK local government, and quite a lot of academic study of its character and consequences, that it ought to be a source of some general propositions of value not only for the UK system itself but for other jurisdictions as well.

2.2 Features of UK local government and why benchmarking matters

The significance of benchmarking in the UK arises out of particular features of UK governance, including the size and scale of local government; its considerable service delivery responsibilities; and its financial dependence on central government.

Vertical fiscal imbalance is a core feature of UK public services, and that drives a great deal of benchmarking behaviour.

Structure and scale

Many areas have one level of principal local authority ('unitaries') and compared to most other jurisdictions they cover relatively large populations, typically between 100 000 and 1 000 000+ residents. This includes most major urban conglomerations, all of Wales and Scotland, and a growing number of largely rural areas. Some county areas retain two tiers of principal local authority — each county authority then contains a number of separately elected district councils. Typically the principal local authority in an area (or the county and districts combined in the two tier areas) will have expenditure responsibilities in the order of £3000 to £4000 per capita. So a unitary authority of 100 000 persons would have a revenue budget of between £300 million and £400 million per annum. If they provide all or most services in-house, the associated staffing requirement will be about 4000 full-time equivalents. About 80 per cent of their expenditure is met by grant from central government, allocated on a needs-based formula. The balance comes from local property taxes (the 'council tax') and from fees and charges.

Tasks

The principal local authorities are responsible for a very wide range of services. These include pre-school provision; primary and secondary education to age 18; and social services and social care for the vulnerable. Local authorities provide social housing either directly or via associated providers. They also have functions for

roads and for transport; economic development; and local regulation. Moreover, since 2000 the principal local authorities have an explicit and a growing role for ‘community leadership’. The precise form varies but this places responsibility on these authorities to plan comprehensively for their areas — including coordination of other local public services (police, fire and emergency, health, and so on) in jointly tackling issues that cross service boundaries such as community safety, or address the needs of groups requiring active collaboration between public services such as the elderly. More recently, the community leadership ambition has extended even further to encompass the role and impact of centrally-run services that are delivered locally, including skills training for example, initiated through a programme called ‘Total Place’ (though since re-labelled by the 2010 coalition government as ‘Community Budgets’).

Central–local relations

Unsurprisingly, given this landscape, central–local relations have long been a significant feature of policy and administration in the UK. They have always been dynamic, and under a degree of tension. In the Thatcher years the focus was on forcing local authorities to accept a degree of marketisation of services, and subjecting them to explicit performance measurement through the agency of the newly created Audit Commission. This continued under the Blair–Brown governments, which further developed an explicit and muscular centrally-run performance regime via ‘arms length’ bodies such as the Audit Commission. The working assumption by central government throughout has been that local government needed to improve and change — although as we shall see in Section Two, that assumption evolved across the period 1998 to 2010.

During the early years of the Blair–Brown governments, both Wales and Scotland gained a significant measure of devolved authority, and thereby gained responsibility for local government generally and for virtually all local government services. For these devolved governments central–local relations figure even more strongly, if only because local government in both nations is responsible for a significantly larger share (about 40 per cent) of their overall expenditure, all of which is financed from the UK level by way of grant aid to the devolved governments (Midwinter 2002). Relationships between central and local governments have tended to be closer in Scotland and Wales than in England. Nonetheless, the devolved governments have also operated performance regimes for local government, again via arms-length bodies including the functional equivalents of the Audit Commission — for example the Wales Audit Office (WAG 2005).

Local government performance and benchmarking

In the UK, local government performance matters well beyond local government itself principally because of the sheer cost and scale of the services they provide, and the level and nature of centrally provided finance and the concomitant limitations and constraints on local fund raising. UK local government effectively inverts the old adage about ‘no taxation without representation’ in that local communities enjoy ‘local representation (and service delivery) without local taxation (even the locally raised and administered council tax has been effectively capped by central diktat)’. For the very most part, local populations do not finance local services other than via the taxes they pay to central government, and there is only an imperfect and not widely understood relationship between the level of local council tax levied and the extent and quality of services provided. For this reason, coupled with the impact of national political sentiment on local elections which are generally fought under mainstream political party banners, local elections provide limited mechanisms for accountability of spending and of service performance.

At the same time, that performance is critical not only to the communities to whom the services are being delivered, but also to many national politicians and government ministers. National programmes are often delivered through and by local authorities, and so national politicians are keen to see them do well and to be able to claim the credit. Local politicians may be just as similarly well motivated but they have in addition the responsibility for actual delivery, with all its attendant problems.

These features of UK local government that prompt extensive performance measurement and associated benchmarking both exemplify and amplify wider tendencies in UK public services. Pollitt (2006) identifies the UK as an outlier in this respect, and traces the cause to a mixture of scale and centralisation, institutions, and culture, an analysis endorsed by Hood (2007) and James and Wilson (2010). The consequences in relation to benchmarking and local government in the UK have been profound. Benchmarking in all its many guises has become ubiquitous, and in so doing perhaps also reached the tipping point at which a ‘logic of escalation’ sets in (Pollitt *et al.* 2010) through which initially few and simple measures become more numerous and comprehensive, and also become summative through league tables and/or targets. These then become linked to incentives and sanctions, with associated pressures for gaming. The measures become more complex, and harder for non-experts to understand, and ownership of the measures becomes more diffuse, all resulting in a decay of public trust. As we shall see, attempts were made in the UK to avert or reverse this logic across the period from 1998 to 2010, but the effort it took to do so is itself testimony to the

underlying ‘natural’ dynamic that Pollitt identifies in public service performance measurement systems.

2.3 Eras and theories of improvement in UK (mainly England) local government

The benchmarking of local government services can be undertaken as an exercise with intrinsic value in order to facilitate ordered comparison and dialogue between service providers, with no more than an eye to avoiding being a laggard in terms of cost or technical standards. But even at this basic service comparison level it is generally stimulated by thoughts of wider progress and learning. When initiated at the governmental level, and especially when initiated by central government with the intention that it should apply to local government, then more is usually afoot, or so it has been in the UK. Benchmarking has been at the heart of a number of attempts by central governments to improve UK local government and its services. Those eras had their prologue in the Thatcher years, and benchmarking had a considerable role then. This was the genesis of the Audit Commission, and the comparisons of service cost that drove the regime of Compulsory Competitive Tendering where local authorities were required to expose (mainly blue collar) services to market forces. It was also when a national set of Performance Indicators was established for local services, for which data were collected and published nationally (Boyne 2002). It gathered force and reached its apogee across 1998 and 2010.

As that momentum increased, it became associated much more explicitly in England for local government with the theories of improvement associated with each of the two eras into which the Blair–Brown years divide of 1998–2005, and 2006–2010. The first was an era of a **central drive for local improvement** and the second was an era of **local drive for self-improvement and self-regulation**. They in turn have given way through a change of central government to a radically different approach, which is best described as a **central drive for localism**. As we shall see, a drive for localism does not mean entirely abandoning a role for benchmarking, though its role is undoubtedly different and considerably reduced, at least in the short term.

1998–2005: Central drive for local improvement

In this period the underlying theory of improvement was that local government would not improve unless it were driven to do so by muscular performance measurement and performance management by the centre. As explicitly articulated

by one of the Blair government's most senior public services adviser, the aim was to drive public services, including local government, from 'awful to adequate' (Barber 2006).

The chosen policy instruments started with the statutory requirement that local authorities should achieve 'Best Value' across all of their service provision, this being the optimal mix of efficiency, quality and effectiveness. This statutory aspiration was linked explicitly to two other instruments. First, local authorities were required to collect data to inform a new set of Best Value Performance Indicators with the data being validated and published in league table form by the Audit Commission. Secondly, they were required to conduct Best Value Reviews of all their services on a rolling programme against a pre-set statutory framework of which comparison with others was a central tenet, with those Reviews also being externally validated by the Audit Commission and the results published (Ball *et al.* 2002).

The Best Value programme was quickly seen as having two major failings. First, the volume of the Reviews themselves was seen to be excessive and indigestible by local authorities and by the Audit Commission alike, and of generally poor quality and lacking the 'bite' of successful challenge to existing poor services. Secondly, it was considered to be insufficiently focused on the things that really matter in determining whether a local authority improved.

The Comprehensive Performance Assessment (CPA) was introduced to remedy these defects, and this became the star performer and the most brutal instrument in the UK local government benchmarking armoury. It involved a periodic, composite, external judgement of each local authority by the Audit Commission of its corporate capacity and leadership, and also its services, with the authority being publicly scored across a range from 'Excellent' to 'Poor'. Those judged to be 'Poor' were then subject to central intervention and support without any choice in the matter.

These new instruments called not only for considerable effort and investment by local authorities, but also significant additional capacity and expenditure by the Audit Commission and by a number of other inspectorates of services (Davis and Martin 2008; Downe 2008; Martin 2006). In due course, CPA evolved during this period into CPA version two, as the 'Harder Test', reflecting the active management and development of the performance regime.

Running alongside CPA were other instruments, including the Beacon Council Scheme, which offered financial and reputational rewards for excellence and which promoted learning between councils from the best practice thereby identified and celebrated (and on which see Section Four, below). This clearly reflected a very

different improvement ‘theory’, but the respective direct costs of CPA and of the Beacon Council Scheme reflect their respective ‘theoretical’ weight in the mix — the expenditure for the Beacons was less than 2 per cent of that for CPA.

2006–2010: Local drive for self-improvement and self-regulation

By 2005 the local government improvement picture had changed considerably. As judged by CPA scores, things had got better generally, and the number of ‘poor’ authorities had declined considerably (Grace and Martin 2008). The Government’s five-year local government strategy of 2005 and the Local Government White Paper of 2006 both signalled that they had responded to the challenges (‘burden’) of a top-down benchmarking approach by shifting from centrally-driven performance measurement to a more ‘voluntary’ (local) system. The new local performance framework remained robust, and aimed to bring together a revised and more outcome focussed set of national performance indicators (approximately 190), local area agreements (LAAs) and comprehensive area assessment (CAA) to provide a coordinated focus for improving services and quality of life for local people. In beginning the move from CPA to CAA, central government underlined the community leadership role of local authorities, and signalled a much bigger role for local government itself in the drive for *self-improvement* and self-regulation. The underlying theory of improvement was that local government was increasingly performing well, and that top-down performance measurement and management both *could* therefore be eased back and, indeed, *needed* to be eased back if further improvement were to be achieved (Barber 2006).

Alongside the shift to a more sector-led approach there was a reduction in the level of inspection, intensifying a trend that had begun to emerge somewhat earlier (Bundred and Grace 2008). At the same time, ‘voluntary’ improvement activity by the Local Government Association and the Improvement and Development Agency intensified and claimed a leadership role (De Groot 2006), focused around the ‘voluntary’ alternative to CPA in the form of the sector-led Local Government Improvement Programme (Jones 2006). Renewed programmatic statements emphasising the need for central government to ‘let go’ and for local government to ‘take responsibility and move beyond compliance’ were published during this period (De Groot 2008).

2010–201?: Central drive for localism

The clamour for ‘localism’ from central and local government alike during the latter years of the Blair–Brown governments could have been a bridge of continuity

across to the new coalition government of 2010. In the event there has been a much greater sense of rupture in that for *ideological* as well as for practical reasons the Coalition government believes that local government should have responsibility for its own improvement as part of a wider theory of the relationships between state, society and public services. ‘Localism’ has been articulated as a major theme, along with idea of the ‘Big Society’ (Tuddenham 2010). The abolition of CAA and of the Audit Commission reflected the new theme, with ‘armchair auditors’ taking centre stage instead of central institutions and frameworks.

It is too early to gauge with confidence how the new framework of ideas and policies will develop, but it is already clear that this new era will not be a stranger to benchmarking. Indeed, the armchair auditors will require consistent, relevant, validated and national benchmarking data to be able to do their job properly. The new ‘shock troops’ are the Efficiency and Reform Group of the Cabinet Office and benchmarking is one of the pillars of their programme (Collier 2010). Further, in specific relation to local government, APSE has already claimed Coalition government endorsement for its current voluntary scheme.

Theories of improvement and the devolved governments

In Scotland and Wales, the instruments of local government benchmarking have been rather different — although to international eyes the similarities may look more significant than the differences. The governments of both Scotland and Wales deployed variants of CPA — Best Value Audit and the Wales Programme for Improvement — that similarly reviewed their local authorities on a ‘whole-organisation’ basis using their principal arms-length public audit body. Both also contained a similar emphasis on the importance of corporate capacity and leadership, and both made provision for intervention by the centre where failure was found. But there were also important differences, notably in the degree to which the results were assembled into scores or published as league tables (Martin *et al.* 2010). Importantly, these differences were expressly associated with different theories of public services change and improvement — against the competitive character and the use of the court of professional and public opinion employed by CPA, it was education and persuasion that was stressed by BVA, and collaboration and consensus by WPI (Martin *et al.* 2010).

2.4 Varieties of benchmarking and UK local Government

Benchmarking in one form or another has featured in all the ‘theories of improvement’ as applied to UK local government. Moreover, each benchmarking instrument carries — at least potentially — a ‘sub-theory’, which helps explain (were it to be articulated) what behaviour it is hoping to stimulate or inhibit through its application. Such theories are not, of course, always made explicit, and if they are they may not be right about the behaviour predicted. Nor is it always the case that where a bundle of instruments are explicitly assembled the resulting composite ‘theory’ will be internally coherent or fully comprehensive. Just as UK governments have been vigorous in their use of benchmarking for local government, so have they also been fairly explicit about what they hoped to achieve and how — but they may not have got it right. There is a quite strong and definite relationship between benchmarking instruments and theories of improvement but it is not always easy to pin down in particular instances.

So this Section reviews a fairly small number of the range of benchmarking techniques applied to local government in the UK with a view to explaining a little more about how they worked and also to explore what seems to work and not work in terms of the types of indicator and the reward/punishment mechanisms deployed, including whether any perverse incentives have been a significant or minor problem. The techniques reviewed are service-based benchmarking clubs; national performance indicator sets; whole organisation assessments; and ‘excellence’ benchmarking.

Service-based benchmarking

There are three initiatives to cover here. First APSE’s benchmarking focuses on the value of services provided by direct labour in local authorities, and can be seen as having its genesis in a partly defensive approach. The data sets have operated from 1998 to the present, and now cover 200+ local authorities. The data sets are practitioner developed in what is entirely a voluntary scheme, and mainly cover ‘manual’ services such as building cleaning and maintenance; civic venues; culture, leisure and sport; education catering; highways maintenance; parks, open spaces and horticultural services; refuse collection; sports and leisure facility management; street cleansing; street lighting; transport operations and vehicle maintenance. The data sets are extensive; for example, there are 30 key performance indicators for building cleaning. APSE see their approach as valuable in the ‘localism’ debate, and celebrate its endorsement by the coalition government (see, generally, www.apse.org.uk).

A second scheme is run by CIPFA, the professional accountancy body centred on the public sector. Their scheme too is voluntary. It has a financial focus, and draws together data on children's services; police; environmental matters (crematoria, cemeteries, waste, etc); housing (rents, homelessness, etc); leisure; personnel matters; and planning. Some of the data sets are long standing (10 years), and some much more recent. Membership costs £600 per local authority per service 'club', and is variable between them (www.cipfa.org.uk) .

A third is operated by the WAO, for example in the policy area of waste collection and disposal. They work to set up benchmarking clubs with the Wales Local Government Association (WLGA) and Welsh Assembly Government (WAG). As the public audit body for Wales with both value for money and performance responsibilities as well as financial audit, the WAO's reasons for operating the clubs are partly to achieve an accurate cost modelling for the review of the national waste strategy and to provide a policy baseline. The benchmarking data allow for greater comparisons between authorities, supporting the sharing of best practice and bringing together service improvement and efficiencies. WAO also provides annual financial reports on the waste management undertaken by local authorities and informs an annual report (WLGA 2010) and a Report to the Wales County Surveyors Group (WAO 2008).

There appears to be little evaluation of the impact or effectiveness of any of these voluntary arrangements, although their continuity and growth clearly indicates that the members appear to get sufficient value to warrant continuing to pay the (albeit modest) annual subscriptions. Nor is it certain that they would be so popular if the wider frameworks of performance measurement were not in place. They are, however, highly consistent with an approach emphasising professional front line leadership of service improvement.

National performance indicator sets

The best example of an instrument in this area is the set of Best Value Performance Indicators (BVPIs) that operated in England from 1998 through 2006 before they gave way to a more outcome focussed National Indicator Set of PIs. The BVPIs consisted of some 200+ indicators across all frontline and corporate services. They were set centrally — albeit after extensive consultation — and were regularly revised and supplemented in light of experience. This had the virtue of ironing out problems of interpretation or data collection and validation, but weakened the continuity of the data sets. The BVPIs were operated by the Audit Commission, and at their height required some 287 pages of guidance to try and ensure that data were collected in a standardised and comparable form (Boyne 2002).

As an example, the BVPI Equality Standard was articulated as the ‘level of the Equality Standard for local government to which the Authority conforms in respect of gender, race and disability’. The purpose of the standard was to provide ‘a framework for delivering continuous improvement in relation to fair employment outcomes and equal access to services to which all Authorities should aspire’. Authorities were required to report the level they had reached according to a six-point scale. This ranged from Level 0 (‘The Authority has not adopted the Equality standard for Local Government’) through Levels 1, 2 and 3 — the latter being that the ‘Authority has completed the equality action planning process, set objectives and targets and established information and monitoring systems to assess progress’ in addition to adopting a comprehensive equality policy (Level 1) and engaging in an impact and needs assessment (Level 2). The highest level (Level 5) is that the authority has achieved targets, reviewed them and set new targets, and is seen as exemplary for its equality programme.

The BVPIs reflect Pollitt’s (2010) ‘logic of escalation’ — not so much in their number, form and character, but inasmuch as they required ever increasing amounts of guidance to enable them to be operated in the preferred manner, and then themselves were put to one side in favour of a revised set of measures which were intended to be more ‘outcome’ oriented. At the same time, however, there was a persistent trend — especially after 2006 — towards reducing the number of indicators which local authorities were required to collect, on grounds that the collection of the data represented a regulatory burden on local authorities.

Whole organisation assessments

The approach of whole-organisation assessment and benchmarking of UK local government had its genesis in the Blair government’s policy concept of ‘Best Value’ and the idea of reviewing specific local government services comprehensively on a rolling programme against a framework of four ‘Cs’ — challenge, compare, consult, and compete — externally validated by the Audit Commission (Ball 2002). That gave way in the face both of the indigestibility of the sheer volume of reviews even by a radically expanded Audit Commission capacity, and the recognition that individual service performance rested as much (especially in the long term) on broader corporate capacity and leadership as it did on the ‘internal’ features of a particular service. The result was a centrally driven programme applying to all local authorities and delivered by the Audit Commission known as CPA. This was the ‘whole-authority’ external assessment against a pre-set framework focussed on leadership/corporate capacity and on key services, leading to a public score and with intervention for low scores, at a programme cost of some £200 million per annum in direct costs.

In Wales and Scotland the very interventionist approach of CPA was explicitly rejected, but both jurisdictions similarly developed on from the initial Best Value methodology to their own versions of whole organisation assessment (Grace 2007; Martin *et al.* 2010). In all three jurisdictions the approach was strongly associated with the development of ‘holistic’ public services inspection (Bundred and Grace 2008).

Excellence benchmarking

Just as significantly, the local government summit body for England also established a ‘voluntary’ whole organisation assessment methodology in the Local Government Improvement programme (LGIP) which evaluated individual local authority performance against an European Framework of Quality Management (EFQM) related model. This had twelve benchmark criteria: forward looking; community leadership; vision; corporate effectiveness; a performance focus; people capacity; values; participatory; member/officer relationships; resource usage; improvement; and well regarded externally. Although voluntary, its underlying approach also received endorsement from government as an improvement method (Bowerman 2002). It used a peer review methodology with mixed review teams, including local politicians from other local authorities, at a cost of some £20 000 per review. Whilst the LGIP was partly a defensive alternative to the government’s Best Value and CPA methods it emphasised a philosophy of ‘improvement from within’ (De Groot 2006), and there is some evidence of positive change associated with the reviews conducted under the LGIP (Jones 2004; 2006).

It was CPA, however, that was widely credited with driving up local government performance between 2001–02 and 2005–06 as judged both by the scores and by the views of almost all the key stakeholders (Grace and Martin 2008; Laffin 2008), although academics have been critical of many of its aspects (Leach 2010; Andrews 2004; Boyne 2004; Martin *et al.* 2010; Jacobs and Goddard 2007; Maclean *et al.* 2007). The public have been less convinced that local authorities have improved (Grace and Martin 2008), although there are data to support a link between public satisfaction with local authority services and CPA scores (Bundred 2006; Bundred and Grace 2008; Ipsos MORI 2007), and the public overwhelmingly support independent external assessment of local authorities.

Certainly CPA appears to have been more effective than the equivalent in Wales (Andrews and Martin 2007), although it is less clear that it has had more impact than the Scottish equivalent (Downe *et al.* 2008).

Insofar as local government has focussed on excellence and best practice identification and exchange, that too has been centrally driven. The Beacon Council Scheme is the most prominent example, with its focus on identifying best practice against pre-set criteria and then spreading the learning to other authorities. Eventually it became a 10-year programme across 1999–2009, funded centrally but operated by a local government summit body, and overseen by an Independent Panel appointed by a government minister. It was initiated partly to offset the widespread impression that the Blair government regarded local authorities as universally in dire need of doing better, although in comparative financial terms it was only a fig leaf of ‘respect’. Originally an England-only programme, Wales eventually established something similar, but rather more ‘collegiate’ in style.

The Scheme invited bids against ten themes each year chosen (after consultation with local government) by the Panel and by government ministries. They were generally fairly complex service/policy themes — for example, child mental health services, regeneration, and asset management — and often had strong ‘local–central’ aspects. It was popular throughout its 10-year life, with typically 250+ applications a year across the 10 themes. It cost some £5 million per annum, mainly in incentive grants to winners, showcasing and knowledge transfer events, specific grants for projects of knowledge transfer such as mentoring of neighbouring local authorities, and administration. Winners were announced at an ‘Awards Dinner’ event, and kept secret till announced.

The Scheme was subject to specific evaluation, and those evaluations were positive overall. They found that participating authorities used the Beacons’ experience to support knowledge acquisition and learning networks, and to underpin knowledge co-creation, and also engaged in actively applying the knowledge acquired to service improvement (Rashman *et al.* 2006). There were also opportunities for ‘vertical’ as well as ‘horizontal’ learning to be taken, which were especially important when so many complex policy problems depend for their success on effective ‘central-to-local policy/delivery’ chains. Such and other benefits do, however, depend upon there being sufficient capacity in the ‘receiving’ local authority (Rashman and Radnor 2005). There is also a broader question about the relationship of Best Practice programmes to wider questions of innovation, highlighting both the growing importance of innovation (Hartley 2005) and doubts as to whether Best Practice approaches do in fact stimulate innovation as compared to ‘top down’ approaches such as CPA. Somewhat counter-intuitively, the latter appears to have the edge (Brannan *et al.* 2008).

A trigger and not a silver bullet

If this overview of benchmarking of UK local government appears somewhat equivocal in the judgement of its efficacy, and that of particular benchmarking instruments, that should not be a surprise. UK local authorities are large, complex, multi-functional organisations that operate in a diverse and turbulent policy environment. As both Hood (2007) and James and Wilson (2010) have observed, there is much still to learn and understand about how ‘performance by numbers’ works and can be made to work better. In contrast to those who see a way forward through focusing on the way benchmarking itself is done (Tillema 2010), the answer if there is one almost certainly lies in appreciating more fully the relationship between a benchmarking instrument and the context of its use (Hill 2006).

An essential part of those varying contexts is the central–local relationship itself at any given time. The range, even at one particular moment, can be considerable (Grace and Martin 2008) and available scenarios as to how that relationship might develop are themselves associated with the deployment of performance instruments within it, and can carry a strongly normative character. Thus, a relationship of ‘targets and terror’ (Coulson 2009; Hood 2005) carries both potential risks and rewards (Jackson 2005) and a potential regulatory burden but one which may pay dividends (Bundred and Grace 2008). Central and local government and regulators alike may remain wedded to such models when they have already passed their optimum effectiveness (De Groot 2008), and when central government needs to let go and local government needs to move beyond mere compliance. In contrast, an era of ‘cooperation and contract’ in central–local relations invites the use of both different instruments and different behaviour (Young 2005) — especially if the focus is switched to achieving desired outcomes rather than merely desirable outputs (Wimbush 2010; Fenna, this volume). If the other end of the spectrum is reached — one that may be characterised as a locally driven approach of ‘initiative and innovation’ — then the role of benchmarking is likely to look very different, and perhaps much less intensive (AC 2007; Moore 2005; Albury 2005). As seen by one of the high priests of public service change and improvement in the UK, it is really a question of whether, for example, the entity to be improved needs to move from ‘awful to adequate’ or is rather at the stage of going from ‘good to great’ (Barber 2006).

2.5 Benchmarking for improvement

In a system like the UK’s, benchmarking has to be part of a suite of improvement tools because its scale and character is significantly subject to the influence of

national government and the consequences of local–central relationships. It has to be seen as an ‘arrow in the quiver’, rather than a silver bullet. At the same time, though, in the UK local government setting benchmarking is subject to counter-productive political and relational influences. These are not aligned to and often undermine the key features that make benchmarking more effective in terms of agreed and stable definitions; reliable data sets and reporting; agreement on the meaning of differences; and using difference as a starting point for inquiry associated with focussed change.

This political context sometimes gives rise to odd policy inversions when it comes to using benchmarking instruments. For example, the devolved governments of Wales and Scotland now operate more centralised performance regimes for local government than does England under the new Coalition government, whereas hitherto they were considerably less so. Moreover, the move to a radically ‘localist’ approach in England may well generate problems with benchmarking that may be similar to those experienced in federal systems — such as ensuring data and definitional consistency and reliability; establishing authoritative and widely supported performance indicators; coping with the extra difficulty of trying to judge outcomes rather than only inputs and outputs; and so on.

A further major contemporary challenge in the UK context, and in many other jurisdictions, will be how to best deploy benchmarking in an age of austerity and the attendant cuts in public expenditure and retrenchment in public services. In the UK at any rate it is widely acknowledged that long-term expenditure reductions will have to draw on change at the tactical, transactional, and transformational levels. At the **tactical** level — tightening efficiency in existing services, shrinking eligibility, and so on — *financial* indicators look to be the most useful. For **transactional** change — improving systems using ‘lean’ methods or better technology, for example — *process* benchmarks are likely to be more relevant. But for **transformational** change — tackling the ‘wicked’ issues, for example, that cross organisational boundaries, or where services are being completely re-designed around customer needs — probably only *excellence* benchmarking will be of any use at all, at least at the initial, innovatory stage when the early adopters are struggling at the leading edge.

Either way, the lesson is fairly clear. It is essential to think about and to deploy a combination of benchmarking (and other) tools from the *improvement* end of the telescope. Adopting an outcome focus, policy makers need to ask themselves: what do you want to get better? What is the current context of change, and what are the key relationships and forces shaping that context? How do you think change will happen — what is your theory of improvement? What will be the role of benchmarking within that? And how best can you optimise that role? This will still

not fashion a silver bullet of change from the benchmarking tools at their disposal, but it will perhaps help to ensure that the triggers for improvement are more likely to work in the right place and in a timely manner.

References

- Albury, D. 2005, 'Fostering Innovation in Public Services', *Public Money & Management*, vol. 25, no. 1, pp. 51–56.
- Andrews, R. 2004, 'Analysing Deprivation and Local Authority Performance', *Public Money & Management*, vol. 24, no. 1, pp. 19–26.
- , Boyne, G.A., Law J. and Walker, R.M. 2005, 'External Constraints and Local Service Standards: the case of Comprehensive Performance Assessment in English Local Government', *Public Administration*, vol. 83, no. 3, pp. 639–56.
- , and Martin, S.J. 2007, 'Has Devolution Improved Public Services? An analysis of the comparative performance of local public services in England and Wales', *Public Money & Management*, vol. 27, no. 2, pp. 149–56.
- Andrews, R. and Martin, S.J. 2010, 'Regional Variations in Public Service Outcomes: the impact of policy divergence in England, Scotland and Wales', *Regional Studies*, vol. 44, no. 8, pp. 919–34.
- AC (Audit Commission) 2007, *Seeing the Light: innovation in public services*, London.
- Ball, A., Broadbent, J. and Moore, C. 2002, 'Best Value and the Control of Local Government' *Public Money & Management*, vol. 22, no. 2, pp. 9–16.
- Barber, M. 2007, *Instruction to Deliver: Tony Blair, the Public Services and the Challenge to Deliver*, Politicos, London.
- 2006, 'Success — it's all in the execution', in Hill, R. (ed.) *In the Matter of How: change and reform in 21st Century public services*, Solace Foundation Imprint, London.
- Bennett, C.J. 1997, 'Understanding Ripple Effects: the cross-national adoption of policy instruments for bureaucratic accountability', *Governance*, vol. 10, no. 3, pp. 213–233.
- Bowerman, M. 2002, 'Isomorphism without Legitimacy?', *Public Money & Management*, vol. 22, no. 2, pp. 47–52.
- Boyne G.A. 2003, 'Sources of Public Service Improvement: a critical review and research agenda', *Journal of Public Administration Research and Theory*, vol. 13, no. 3, pp. 367–94.

-
- 2002, ‘Concepts and Indicators of Local Authority Performance’, *Public Money and Management*, vol. 22, no. 2, pp. 17–24.
- and Enticott, G. 2004, ‘Are the “Poor” Different?’, *Public Money & Management* vol. 24, no. 1, pp. 11–18.
- Bradbury, J. 2003, ‘The Political Dynamics of Sub-State Regionalisation: a neo-functional perspective and the case of devolution in the UK’, *British Journal of Politics and International Relations*, vol. 5, no. 4, pp. 543–75.
- Brannan, T., Durose, C., John, P. & Wolman, H. 2008, ‘Assessing Best Practice as a Means of Innovation’, *Local Government Studies*, vol. 34, no. 1, pp. 23–38.
- Broadbent, J. & Laughlin, R. 1997, ‘Evaluating the “New Public Management” reforms in the UK: a constitutional possibility’, *Public Administration*, vol. 75, no. 3, pp. 487–507.
- Bundred, S. 2006, ‘The Future of Regulation in the Public Sector’, *Public Money & Management*, vol. 26, no. 3, pp. 180–88.
- and Grace, C. 2008, ‘Holistic Public Services Inspection’, in Davis, H. and Martin, S.J. (eds), *Public Services Inspection*, Jessica Kingsley, London.
- Clarke, J., Gewirtz, S., Hughes, G. and Humphrey, J. 2000, ‘Guarding the Public Interest? Auditing public services’, in Clarke, J., Gewirtz, S. & McLaughlin, E. (eds), *New Managerialism, New Welfare?* SAGE, London.
- Collier, S. 2010, ‘The Work Programme of the Efficiency and Reform Group’, presentation to the *Efficiency in Central Government Conference*, 30th November 2010, London: Inside Government.
- Coulson, A. 2009, ‘Targets and Terror: government by performance indicators’, *Local Government Studies*, vol. 35, no. 2, pp. 271–81.
- Davis, H. and Martin, S.J. (eds) 2008, *Public Services Inspection*, Jessica Kingsley, London.
- DCLG (2011) *Future of Local Public Audit: consultation*, Department for Communities and Local Government, London.
- De Groot, L. 2006, ‘Generating Improvement from Within’, in Martin, S.J. (ed.) *Public Service Improvement*, Routledge, Abingdon.
- and Rogers, P. (eds) 2008, *Getting so Much Better all the Time*, London: Solace Foundation Imprint.
- 2008, ‘Challenges to the Approach to Improvement’, in Grace, C. and Martin, S.J. (eds) *Getting Better all the Time?*, IDEA, London.

-
- Downe, J.D., Grace, C.L., Martin, S.J. and Nutley, S.M. 2007, 'Comparing for Performance Regimes in England, Scotland and Wales', *ESRC Public Services Programme Discussion Paper 0705*, www.publicservices.ac.uk/library.
- 2008, 'Inspection of Local Government Services', in Davis, H.D. and Martin, S.J. (eds), *Public Services Inspection*, Jessica Kingsley, London.
- , Grace, C.L., Martin, S.J. and Nutley, S.M. 2008, 'Best Value Audits in Scotland: winning without scoring?', *Public Money & Management*, vol. 28, no. 2, pp. 77–84.
- Grace, C. (ed.) 2007, *Comparing for Improvement*, Solace Foundation Imprint, London.
- and Martin, S.J. (eds) 2008, *Getting Better all the Time?*, IDeA, London.
- Hartley, J. 2005, 'Innovation in Government and Public Services', *Public Money & Management*, vol. 25, no. 1, pp. 27–34.
- Hill, R. (ed.) 2006, *In the Matter of How: change and reform in 21st Century public services*, Solace Foundation Imprint, London.