30 years down the wrong rabbit hole: how we got there and how we get out

Richard Hamblin
Leadership Development Centre 2018 Fellow
July 2018

Acknowledgements and disclaimer

What follows is an ordering of my thinking about *why* the public sector needs to change its relationship to data. It's essentially a combination of reasonably large amounts of reading and a reasonably large amount of experience in trying to use information to make health services better. This was undertaken as a starting point for my Leadership Development Centre fellowship which was much more concerned about *how* the public sector could change its relationship to data. As such doing this was only possible because of LDC's generosity in awarding me a fellowship and that of my employer, the Health Quality and Safety Commission, in allowing me study leave to read, think and organise my thinking. It's what I thought as I embarked on this process of learning. It's probably contentious, certainly challengeable, possibly wrong, and likely to change as I study more. It also certainly does not necessarily reflect the views of either the LDC or the Commission.

Introduction

It's all very well arguing that we need to use data in a different way and therefore we need data teams that look like *this* and are led like *that*, (and by the end of this fellowship process I may have some idea of what this and that look like), but a sceptic could perfectly reasonably ask "why do we need to use data in a different way?" What follows is my answer, the reason why I embarked on this exercise.

Mansell's challenge

A fascinating 2015 paper by James Mansell for the Productivity Commission, *Handing back the social commons*, critiques the public social sector along the following lines. "It is not value focused, it is slow to innovate and does not reward success well."[1]

In reflecting on why this should be he identifies "30 years of Taylorist thinking" as a major, if not *the*, cause. Frederick Winslow Taylor, the American engineer, is often identified as the father of scientific management. His concern for understanding and optimising workflows to maximise labour productivity is often associated with development of mass production in the first part of the 20th century[2].

In addition to the atomised, "assembly- line" deconstruction of task, this approach depends upon measurement, so for those who are data professionals, if Taylorism is dominant paradigm for public sector management, measurement in its interests will become the dominant mode of work.

Whatever the virtues of Taylorism in its concern for understanding process (and as such it is an obvious forerunner to the CQI philosophy which is widely venerated), eliminating waste and unwarranted variation, and controlling costs, organising public services around these principles has, as Mansell correctly identifies, had considerable problems. There is more interest in reducing costs rather than increasing effectiveness, structural incentives tend to drive incremental improvements rather than radical shifts, and silo thinking, cherry picking and cost shifting abound.

The potential risks of imposing such a reductionist view of work as Taylor's, designed for the efficient operation of a production line, to the complexities of public and social service should be so obvious that its ubiquity in public services needs more explanation.

Of knights, knaves and Henry Ford

Mansell's comment on all this, "Henry Ford and the 1900s demonstrated the value proposition of doing this… the public sector caught up in the 1980s", would seem to imply that the use of these approaches was simply a technocratic decision about the most effective way to organise themselves that the public sector made after, presumably, being asleep for the best part of century. This is, I think, to misinterpret history.

[1] Mansell JH (2015). Handing back the social commons. New Zealand Productivity Commission

[2] Taylor is often closely associated with Henry Ford's development of the moving production line, although the links between them are increasingly disputed. Certainly Ford seems constitutionally unable to share credit with anyone else.

Throughout the English-speaking world (and to a lesser extent, western Europe) the 1980s were marked by profound debate about the role, limits and *ethos* of the state. A serious articulation of the change is in Le Grand's Knights and Knaves[3] [4]. This states that the prevailing assumption behind welfare states in the post second world war period was that the motivation of those working within them had "knightly", altruistic motivations. Thus, management in a meaningful sense of the word was unnecessary as those involved would "do the right thing" by those that they served. Given that public sector workers are typically around 20 per cent of the workforce, the base assumption of their universal altruism is questionable. Even if it were not, the corollary that they therefore need not be managed is clearly nonsense. To give one example from health, given finite resources the needs of the individual patient versus those of the population are always in tension. To place upon an individual clinician the responsibility to decide the trade-off for each interaction with a patient and expect justice to result is unrealistic, not to say cruel, on all concerned.

As these tensions became increasingly acute in the 1980s, the "knightly" thesis was challenged. Public sector workers were instead "knaves" motivated by self-interest. In as far as this goes it is a useful corrective to starry-eyed assumptions, but again its development and corollary create problems. The development is that not only are public sector workers self-interested, but they are cynically exploiting the public by hiding in systems where there is no meaningful accountability, no risk of lost market share and income, no threat to tenure. Public servants are lazy wasters who couldn't hack it in the real world, doing as little as possible[5].

The assumption of inevitable public sector knavery is frankly as nonsensical as the presumption of saintliness that it replaced. However, it's interesting that the contempt towards public sector workers implicit in such a view mimics Taylor's towards assembly line workers *"Now one of the very first requirements for a man who is fit to handle pig iron as a regular occupation is that he shall be so stupid and so phlegmatic that he more nearly resembles in his mental make-up the ox than any other type[6]."*

In fact, it would be unfair to characterise the entire approach to managing the public sector as "Taylorist". The concept of New Public Management as defined by Dunleavy[7] and others as consisting of disaggregation, competition and incentivization was far broader in its scope. Yet the heart of Taylor's approach to measurement, that it is a mechanism for control (or at least judgement), is fundamental for this model to be implemented. Fascinatingly, Dunleavy

[3] Perhaps most fully expressed in Le Grand J (2003), Motivation, Agency and Public Policy: Of Knights and Knaves, Pawns and Queens Oxford University Press, New York

[4] Another explanation is that this view was given credence by the BBC situation comedy *Yes Minister* in which self-serving civil servants were able to run rings round a caricature of a craven government minister. Mistaking a light-hearted comedy of manners for serious satire is forgivable; mistaking it for *documentary* rather less so…

[5] This is not hyperbole. Even thirty years on, it is a common trope of the political right See https://www.businessinsider.com.au/britains-firing-every-lazy-government-worker-2012-5?r=US&IR=T https://www.adamsmith.org/blog/politics-government/lazy-civil-servants

[6] It's worth saying that Taylor did still advocate raising living standards for workers- something Ford actually did

[7] Dunleavy P, Margetts H, Castow S, Tinkler J (2005) New Public Management is Dead – Long live digital-era governance, Journal of Public Administration Research and Theory, Oxford.

and colleagues identify measurement and publication to have been one of the few parts of the trend that was still going forward by 2005 (although the extent to which these had mutated into what Bevan and Hood dubbed "targets and terror" is interesting to consider).

The point of this is to recognise that how we have treated data for the last 30 years was not a neutral technocratic choice. The public sector did not "discover" Taylorism; for good or ill it was imposed upon it for ideological reasons[8]. Yet ironically, whatever the rights and wrongs of the ideology, there are good technical reasons why the Taylor inspired approach to measurement has proved problematic.

Perhaps the greatest irony about Mansell's characterisation is New Public Sector Management ideas, which foisted Taylorism onto the public sector were inspired (or at least justified) precisely by the sense that the Public Sector was "not value focused, it is slow to innovate and does not reward success well."

Why the Taylor doesn't fit

As the quote above demonstrates, Taylor viewed much of humanity with contempt. Given he came from a time when racial eugenics was not only acceptable but fashionable, holding such views was unexceptional, but they are surely antithetical to the values commonly held in the public sectors of developed nations. However incoherently held, clumsily expressed or ill-thought through the view that all people are worthy of dignity, respect and value is an underpinning aspiration of public service operation[9]. The importance of a fit of values between sectors is perhaps underestimated[10].

Further, the lack of trust implicit in centralised performance monitoring of those concerned with actually delivering services may be counter-productive. A long-term critic of this approach, John Seddon, charges that it in effect assumes that an infallible centre is better able to predict and describe the demands on public services than those charged with leading their delivery at a local level[11], and this leads to disengaged public sector staff, more expensive services and poorer outcomes for citizens.

However, even if we ignore these ideological problems, there are good technical reasons why the performance measurement approaches that underpin New Public Management (NPM) it may not be appropriate. Osborne and colleagues[12] emphasise that a fundamental weakness of much of the thinking about behind NPM is that it draws lessons from manufacturing rather than services. This is, I think an important distinction. Services are distinct from products being "intangible, process driven and based upon a promise of what is to be delivered". One example of this is the role of the consumer of a (manufactured) product and a consumer of a service. The former is passive while the

[8] And the people doing the imposing had an ambivalent if not antagonistic view of the imposed upon.
[9] The point that services frequently fail to deliver anything like this is, here, moot. The issue is of aspiration and self-perception.
[10] Would it be possible to take techniques of predictive analytics developed by the gambling and pornography industries and simply apply them to the provision of healthcare? Does the argument that maths is just maths really hold up?

[11] Seddon, J. "The Whitehall Effect: How Whitehall became the enemy of great public services and what we can do about it" Triachy Press. 2014.

[12] Osborne S, Radnor Z and Nasi G (2012) A New theory for public service management? Toward a (public) service –dominant approach, *American Review of Public Administration*

latter is involved as a shaper, co-producer and evaluator of the service as it is delivered. Other assumptions that production and consumption are different processes with different costs that apply to manufacturing of goods do not apply to provision of a service. All this means that assumptions about the benefits of standardisation of process, reduction of labour costs and other shibboleths of scientific management may not apply to the provision of services, especially public services which in an ideal world the consumer might prefer not to consume!

Another problem of applying techniques from manufacturing to public services lies in the essential difference between complicated and complex. While manufacturing may be extremely complicated in the number of processes and products that need to be brought together, they are inherently predictable, and the feedback loops inside the process are limited. Provision of services, especially ones which may well be unwanted and which are part of a suite of services which may conflict with each other is much more prone to the effects of feedback within the system. A change at point A may have completely unforeseen circumstances at point B. Yet the effects may depend upon whether circumstance C, D or E pertains at point F. And so on.

This has profound significance for measurement as we will discuss in the next section.

One final thought is that the Taylorist contempt for the worker in all this increasingly being shown to be problematic. The move from triple to quadruple aim in healthcare, explicitly recognising the centrality of ensuring that the workforce is sufficient in supported if the other aims of outcome, experience and system effectiveness is to be achieved challenges both Taylor's mechanistic mindset and the NPM assumptions of knavery.

By 2005 Dunleavy and colleagues judged that many of the more explicitly market oriented elements of NPM had stalled or were in retreat. Despite a general shift to centre right governments in English speaking and western European countries since then, this pattern has continued. However, the hard-edged performance measurement that it begat remained, metastasizing into regimes of "targets and terror."[13]

Soviet nails and synecdoche

"the point about Faustian pacts is that you always get precisely what you asked for you and precisely what you didn't want" – Terry Pratchett

"Once upon a time, there was a factory in the Soviet Union that made nails. Unfortunately, Moscow set quotas on their nail production, and they began working to meet the quotas as described, rather than doing anything useful. When they set quotas by quantity, they churned out hundreds of thousands of tiny, useless nails. When Moscow realized this was not useful and set a quota by weight instead, they started building big, heavy railroad spike-type nails that weighed a pound each."

It's a great story, although almost certainly an urban legend built on a 1950s cartoon (presumably produced under Kruschev's brief proto-*glasnost*), but it hints at a truth. When we set a target as a measure by which all effort is judged, all effort will go in meeting the target in the most efficient way possible. And from here education gives way to examination practice, "hello nurses" ensure ED triage in 5 minutes and so forth.

The problem is that for the target to be measurable, achievable and attributable it has to be refined away from its intention. Thus "all patients should be treated in a timely manner" becomes "no one

[13] Bevan G and Hood (2005) Governance by Target and Terror: synecdoche, gaming and audit, Westminster Economic Form

should wait more than 6 hours in ED" and "we should treat x elective patients". Similarly, "we should keep children healthy" becomes, "we should reduce emergency admissions for conditions sensitive to ambulatory care for people aged 0-14".

This approach has a habit of breaking down. Smith[14] identifies at least eight potential unintended consequences including management distraction and gaming. Further as Bevan and Hood (ibid) recognised, a system built around such measures had to rely on an assumption of synecdoche – that the part could represent the whole. For example, a health service would be good if it made sure people got seen quickly enough. The risk with this is always that other important factors (i.e. that once seen the treatment would be effective, acceptable to the patient and delivered humanely) get ignored.

Even worse, frequently the achievement of the target ends up being decoupled from the aim that it is supposed to achieve. Sometimes referred to as Goodheart's law, it is elegantly described thus[15]:

- Superiors want an undefined goal G.
- They formulate G* which is not G, but until now in usual practice, G and G* have correlated.
- Subordinates are given the target G*.
- The well-intentioned subordinate may recognise G and suggest G** as a substitute, but such people are relatively few and far inbetween. Most people try to achieve G*.
- As time goes on, every means of achieving G* is sought.
- Remember that G* was formulated precisely because it is simple and more explicit than G. Hence, the persons, processes and organizations which aim at maximising G* achieve competitive advantage over those trying to juggle both G* and G.
- P(G|G*) reduces with time and after a point, the correlation completely breaks down.

Hence "hitting the target and missing the point".

However, the argument to just abandon targets doesn't stack up. Until the point where G and G* become uncoupled, benefits compared to relying on "knightly" motives accrue. One third of English ambulance services cheated to hit a response time target, but nearly all improved their actual performance in the process. In contrast the Welsh ambulance service's performance reporting was unsullied, but performance remained unchanged, half of all patients waited too long for the ambulance to turn up.[16]

The key therefore is to avoid this uncoupling, and this requires the centre to use measures which do not over-specify what localities need to achieve, and instead allow them to show a/ what they are doing to achieve the overall goal, and b/ how this is going to work. The recent abandonment of New Zealand's Better Public Service Targets[17] is, I believe, broadly wise. These were largely examples of measure susceptible to the Goodheart's law.

[14] Smith P (1995) On the unintended consequences of publishing performance data in the public sector, Journal of the Royal Statistical Society

[15] http://lesswrong.com/lw/1ws/the_importance_of_goodharts_law/

[16] Bevan G and Hamblin R (2009) Hitting and missing targets by ambulance services for emergency calls: effects of different systems of performance measurement within the UK, Journal of the Royal Statistical Society.

[17] http://www.ssc.govt.nz/better-public-services accessed 10 July 2018

The CEO survival kit, guerrilla warfare and Saturn's children

Just under twenty years ago I was taking the CEO of an NHS trust through a set of quality measures related to his hospital. While their financial management was excellent and they had a good track record of meeting various waiting time targets, some specialty specific mortality and readmission rates raised enough questions to warrant further consideration. His response was interesting. He had, he said, a CEO survival kit which consisted essentially of meeting his waiting targets and staying in the black. While quality measures were interesting to consider, no-one was going to fire him over them. So, this was interesting, he'd share it with his medical director and they would look further at it, but my suggestion that their routine performance reporting should move beyond achieving targets into the quality of care provided was naïve. He was probably correct.

This wasn't a bad man or a bad leader, far from it on both counts. But he was responding to the incentives that had been placed into the system. This was Goodheart's law in action.

In the two decades since there has been some improvement. The use of the Hospital Standardised Mortality Ratio (HSMR) as a single summary measure of quality, while problematic in its interpretation, at least stimulated a recognition that outcomes of care mattered. Sadly, what it stimulated was often manipulation of recording practice (for example of which cases were palliative care) as a mechanism to adjust results. Similarly, increased publication of performance measures, some surgeon level outcomes reports and scorecards, patient experience surveys and patient safety indicators, Atlases of healthcare have changed the landscape significantly. Quality now matters. HQSC has, in New Zealand, contributed to this and in its approaches to measuring unwarranted variation through its Atlas and patient safety through Quality and Safety Markers have been innovative and have, at least at the margins, been associated with apparent improvements.

Yet even as we see these improvements we have a growing sense of unease that this approach has stimulated improvements only in the areas that are measured, rather than stimulating the sort of radical change that would create wide improvement. The CEO's survival kit may have extended its scope, but it remains in place.

The use of the word "survival" is telling. These sorts of regimes are characterised above all by a lack of trust. This means that the technical solutions proposed to address the effects of Goodheart's law (using balancing measures to pick up perverse consequences[18], not defining the measures used prior to their application (in the manner of the relationship of an exam to a curriculum[19]), or linking the targeted processes to the outcomes they are supposed to achieve[20]), can only mitigate, and not avoid the problem. The lack of trust continues, and with it the issue of central imposition of very precise actions (when they have no actual basis in evidence) with no regard to local conditions creates cultures and behaviours (such as the CEO survival kit) which are unlikely to improve public services.

I believe that there are ways out of this which will be discussed below (and the New Zealand health sector is making tentative steps in this direction). Here however I am more concerned about the pernicious effects of this type of regime on our attitude to data and specifically on data teams.

---

[18] Toma M, Dreischulte T, Gray N et al (2017) Balancing measures or a balanced accounting of improvement impact: a qualitative analysis of individual and focus group interviews with improvement experts in Scotland, *BMJ Quality and Safety*

[19] Bevan G and Hood C (2004) Targets, inspections, and transparency *BMJ*

[20] Hamblin R, Bohm G, Gerard C et al (2015) The measurement of New Zealand healthcare, NZMJ

The consequence of using measurement as a mechanism for control is that those on its receiving end learn to adapt their behaviour to minimise the offence that this perceived control creates. The techniques for doing this are easy for a halfway competent manager. If you use data from period x to show poor performance, I will use data from period x+1 to show that is improving. If you use a particular indicator that makes me look bad, I'll argue that the indicator is wrong and calculate the one that makes me look good. If you don't standardise the results, I will. If you standardise the results for age, I'll add sex, ethnicity and socio-economic status. And so on. What follows is a fight for power over data. Rather than using data as a prompt for inquiry and action, seeking to understand causes and working as allies to understand and improve, we spend our analytic resource re-litigating the data. In fact, we have built an entire cottage industry around this and use our data teams to support this and little else. Our data teams are thrown into the trenches of a conflict: the centre with nominal firepower on its side, the outgunned local providers fighting a guerrilla battle back. What is lost is the opportunity to use data as a prompt for action, reflection, appreciative enquiry and learning.

My experience is that the public sector has a habit of recruiting bright graduates, often from maths, stats and the "hard" sciences and then wasting their talents to produce a series of key performance indicator reports based upon a pre-agreed template, running pre-written code, using a pre-ordained list of data quality checks (which have little relevance to whether the measure reflects reality). These simple data processing jobs are dull, repetitive and frequently, because the measures have not been developed with people who actually deliver or received services, provide no insight to those involved in their delivery. These measures are then reported in ways (typically R/A/G status- borrowed from project management disciplines) which are inappropriate to provide any support to delivering ongoing good services, or improving those that need to.

The consequence of this is shown in recruitment. Having done this for twenty years in two different systems and at local and national level, two things have been consistent. It is relatively straightforward to recruit extremely good graduates into the public sector; it is extremely difficult to recruit people of such high quality with five years' experience into career grade posts. Essentially in those first five years of a career what we do either encourages people to leave or knocks out every bit of innovation, imagination and delight out of them. Analysts are either "lost" or "squashed". Like Saturn, we have a tendency to devour our young[21].

In our present situation this is tragic.

Complexity everywhere we look

In the developed world, public services have an entirely new opportunity to use data to address the considerable challenges they face. In order to seize this opportunity, we will need to keep, cultivate and develop the analytic community.

The increasing complexity of what all public services are dealing with is widely acknowledged. Health and social care is dealing with an ageing population with complex overlapping health needs, where the goal is increasingly about management and optimisation of quality of life rather than cure. More can be done, but resources are not increasing at a pace commensurate with innovation. At the same time issues related to the pressures of living from obesity to mental distress are becoming more urgent. In education, shifting patterns of employment and an increasingly globalised economy are raising challenges about what the purpose and success criteria of education actually are. Other pressures for New Zealand include rapid increases in incarceration in the last five

[21] https://en.wikipedia.org/wiki/Saturn%27s_Children

years, increasing demands upon transport infrastructure and availability of affordable high quality housing. All of these operate against a backdrop of ongoing fiscal tightness, and all interact with and affect each other.

There is no space here to raise the issues of "wicked problems", complexity theory, the challenges of increasingly interconnected populations, the rise of ethno-nationalism and populism and the limits of growth against the capacity of the environment to absorb it. Suffice to say that conceiving the challenges of the public sector in managerial and consumerist terms as was common in the 1980s and 1990s (just be more customer focused/ goal oriented/ efficient) is unlikely to meet these concerns.

In this environment, one cause for optimism is the increasing capacity for rigorous data analysis to spot, predict, diagnose and suggest solutions to problems. This comes from the development of the idea of what is often called 'big data' but which probably better described as 'discovery analytics'[22]. This may be described as the combination of a range of techniques which have become possible through the increase of computing power. They include: the linkage of large amounts of different data; identification of patterns and relationships across linked data sets; the capacity (though controversial) of identifying patterns and addressing these; approaches to mining unstructured data to spot risks more effectively and quickly[23]; and the ability of machines to "learn" – refine their algorithms through their repeated application. To this I would add the development of visualisation software to allow managers to understand and investigate the situation that pertains to their service rapidly and intuitively.

New Zealand has, perhaps uniquely, invested in an infrastructure (the Integrated Data Infrastructure) which should, in theory, allow the public sector (or anyone else with access to the data) to start using these new approaches to design, implement and evaluate novel solutions to public challenges.

The question then arises as to who should have access to this sort of data. My sense is that there is the frustration at the progress made in the use of the asset by the public sector has led to siren voices advocating opening access wider. My view is that this is problematic. The assumption that there are an army of analysts sitting in their bedrooms, waiting to run open source machine learning-software over the data and willing to act *pro bono* is frankly a fantasy[24]. What opening up would mean is effectively to make the data available to private corporations who would of course seek to monetize the asset. This may not be intrinsically bad, but in light of the current revelations about Cambridge Analytica[25] such an opening up in the near future, unless issues of governance and social licence inherent in this approach are worked out and agreed, this strikes me as a profoundly dangerous for governments to follow.

---

[22] Russum P (2011) Big Data Analytics, *TDWI Best Practice report*

[23] Griffiths A, and Leaver M (2018): Wisdom of patients: predicting the quality of care using aggregated patient feedback, *BMJ Quality and Safety*

[24] The comparison with the precipitate downsizing of local government auditing in the UK, justified on the grounds that making local authority data publicly available would stimulate "armchair auditors" is telling. In London alone this has been followed by $100m spent on planning an unnecessary bridge that never got build, the off-loading of the Olympic Stadium to a group of pornographers at massive disadvantage to the public purse, and Grenfell Tower. The success of the policy is somewhat open to question.

[25] https://www.theguardian.com/news/2018/mar/18/what-is-cambridge-analytica-firm-at-centre-of-facebook-data-breach

However, unless we create a broader and bigger cadre of analysts who have the skills, time and mandate to use the data in this way, the clamour for a wider opening up of public data will become unanswerable. We have comparatively little time to respond to this, and in my view a move away from KPI culture and revision of the role of analytic departments within ministries and other central agencies will be essential.

The threefold role of the analyst

So how should this be done? Inside central agencies we need to build and train analytical teams to play three roles.

- Research, in the traditional sense of the word;
- Discovery analytics using new machine learning and other techniques to mine existing data sets and feeds to gain insights;
- Reporting – we still need to do this, but we need to do it in a very different way.

Teams working in data should undertake work using all three roles.

Recruitment should be based on a mix of skills (Wellington is blessed with a range of post-graduate courses which provide the skill sets we need). There are a range of implications around resourcing, training, management, incentivization and so forth that need working through (which this fellowship is designed to define).

The following section defines each of these roles in greater detail. The next section covers my proposal for how performance reporting for services needs to change radically away from centrally controlled Taylorism to locally specific, mutually agreed progress towards more broadly defined social goals, with central reporting on this being conceived a support for local teams rather than a traditional accountability mechanism.

Research

By this I mean collection (or harvesting), analysis and presentation of data to answer a specific, pre-defined question. This is specifically about the testing of pre-defined hypotheses, uses inferential statistical tests to understand difference, change and relationships. The approach is linear and planned. There is no "discovery" element to this role (indeed such an approach would rightly in this context be referred to as data dredging). The basic skills required, those of defining measurements, collecting and analysing the data, statistical modelling and testing should be bread and butter for a well-trained analyst.

This role has value in spreading generalizable knowledge. As such it can support service improvement or policy development by identifying and demonstrating best practice. The work however is often quite long-term and discrete each time, with approaches hard to reuse. As such the work has to be of great enough significance to justify the investment of time.

Discovery analytics

Sometimes referred to as 'big data analytics', this refers to exploiting the availability of large linked data sets to identify relationships and draw out and test hypotheses which would not be immediately intuitive. This has a number of ethical dangers and risks the pursuit of spurious correlations but also great promise in a number of areas. The social investment approach of targeting at risk individuals and intervening early to avoid adverse outcomes is perhaps the most discussed use of these approaches, there are however, other less contentious uses to which the approaches can be put. These include:

- Early identification of systems under strain. The discovery analytics approaches can be used to identify common leading risk factors for organisations, services or teams where significant failure has occurred. These can then be used to predict where systems are at risk and early intervention can be undertaken. This is a step on from approaches which spot outcomes going out of control and seek to respond to that. The approach here is to spot and identify risk rather than act after the fact.
- Policy development and testing of intervention logics. The development of intervention logics as a way to design policy and services provides a framework that is useful to clarify thought. Yet usually the various drivers listed and their relationships with outcome are based on expertise and assumption rather than data. Discovery analytics provides both the opportunity to demonstrate the assumed relationships, and test that they still hold when changes are made to the system.

I believe that these two roles should form the majority of the workload for analysts working in the public sector. In order to create the space to do this, we need to rethink radically how we do performance reporting.

A better way to report

The evidence for targets, accountability arrangements and so forth is at best equivocal. By narrow definition things are achieved, but at considerable cost (straightforwardly fiscal, opportunity, morale and ethos). The victories easily become pyrrhic. The question follows whether this is a natural corollary of targets/performance reporting/accountability regimes, or whether this is *how* these have been introduced. My belief is the latter. The consideration of measurement in the publication sector made throughout this paper can summarised as follows: accountability has been conceptualised only in adversarial, low-trust, terms – measures are always dials to judge not tin-openers to question. In such a system, targets have become a mechanism for the centre to second guess and countermand local management. Their bad name is in such a context understandable.

Yet in a representative democracy it cannot be wrong for government, chosen by voters and spending their money to ask tax funded services to achieve outcomes that voters have indicated, through their votes and other means, that they desire and are at least grudgingly willing to pay for. How do we resolve this tension?

The answer lies I think in reconceiving the level at which targets operate. "National targets" should operate at a high conceptual level linked to what a service incontrovertibly designed to achieve. So for healthcare this is relatively uncontroversial, essentially some variant of people living longer, healthier lives. This is a concept with relatively robustly defined (if not straightforwardly calculated!) measures like Disablity adjusted life years (DALY) lost. For other services such as education and justice both the purpose and the overarching measure may need more thought.

The next step is one of embracing subsidiarity. The problem with target regimes is that they have sought to tell local services precisely what to achieve and how, with no regard to local variance, issues, culture or demands. Setting the target at a higher level allows a more thoughtful conversation with the local service about what the local priorities to achieve the desired goal are, and what needs to be done to progress this. Ideally the local service should be working with its local population to define (co-produce in the jargon)what change should be made and how it should be measured. In order to balance these local defined measures and targets measures that

This does not mean an end of accountability. On the contrary. The local service may define what it needs to do to achieve improvement, but it is held to account both for doing what is says it intends

to, and that it selected the right thing to do to make. An agreed process change which doesn't deliver the change in outcomes envisioned needs to be reconsidered.

This approach assumes that local services both have a desire to get the best results possible for the populations that they serve, and have the knowledge of their local situation to identify what is most likely to work. What then is the role of analytic teams at the centre?

At heart I believe it to be about support tempered by oversight, rather than oversight per se. In practical terms this means four things.

- Measurement governance – determining data sources, approaches to developing measures and indicators, valid approaches for testing difference and change and so forth
- Provision of research and discovery analytics to support localities to find solutions and build and monitor intervention logics
- Smart reporting – data visualisation and information for improvement rather than league tables (filterable dashboards that allow managers to dive into the data to identify anomalies to understand and address would be a good example of this. Presentation of data in ways that allows perverse effects to be identified and understood would be another example)
- Public reporting in ways that are genuinely used and understood by the public.

Interestingly, the health sector has, through the System Level Measures Framework, actually moved in this direction. Conceptually, the approach of high level measures which local alliances work together to address seems right. The attempted process in involving the sector has been an appropriate one. Taking this as a model for how measurement is used within the sector, providing more resource and priority would be an excellent move. We have made tentative steps in this direction. Let's take several more.

Replacing existing target and performance reporting regimes with this model would mean an end to the infrastructure of guerrilla warfare where the centre and sector were primed to argue with each other about the meaning of every figure, where analysis that "fed the beast" of arbitrary targets but did nothing to improve outcomes for people was abandoned. A corollary of this is that manual reporting of aggregated data would be replaced with automated reporting using data that all agreed was an authorised version of reality, using metrics that all agreed measured stuff that mattered in ways that made sense. The sector would never complain that they could not replicate the centre's analysis because there would be sufficient trust in the data and measures that there would be no need to undertake this. Both the sector and centre would free their analytic resource to spot and understand issues as they emerged, the former at an operational level and the latter at a policy level. And time would be freed up for the crucial roles of research and discovery analytics that will be essential to address the profound and complex issues in public service that confront us.

…and curiosity abounds

In 2013 in response to one of the English NHS's regular scandals of poor quality services, the leading health care quality guru Don Berwick was invited to look at what could be done to prevent a recurrence. It is probably fair to say that his prescription was not what his commissioners were expecting or were looking for. Rather than advocating more regulation, more competition and the like, his response concluded with a letter to NHS staff which envisaged a culture where,

"measurement is not a threat, it is a resource; ambition is not stressful, it is exciting; defects are seen as opportunities to learn; and curiosity abounds".[26]

This seems to me as fine a prescription for twenty first century data team as any other. The trick now is to define, build, lead and sustain them.

[26]https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/226703/Berwick_Report.pdf