New development: The paradox of outcomes—the more we measure, the less we understand

Toby Lowe


To cite this article: Toby Lowe (2013): New development: The paradox of outcomes—the more we measure, the less we understand, Public Money & Management, 33:3, 213-216

To link to this article: http://dx.doi.org/10.1080/09540962.2013.785707
Outcomes-based performance management (OBPM) frameworks have come to dominate social policy thinking across the world (Perrin, 2006). They come with a variety of labels, from ‘outcomes-based evaluation’ (Schalock, 2001), ‘results or outcomes-based accountability’, (Friedman, 2009), ‘results based management’ (Mayne, 2007) or simply ‘performance management’ (Wimbush, 2011). The most recent form is ‘payment by results’. All these frameworks posit that the effectiveness of any social policy intervention should be judged by measuring its impact against outcome indicators—that is measures of the impact that interventions make in the lives of those affected by them.

Yet literature around OBPM is replete with evidence about how difficult it is to undertake effectively. Studies have demonstrated problems concerning how to define indicators for outcomes that are meaningful but that do not distort practice, how to allocate responsibility for shared outcomes, the distorting effects of performance management, and unintended consequences. (See, for example, Perrin, 1998; van Thiel and Leeuw, 2002; Bevan and Hood, 2006; Wimbush, 2011; and Sager et al., 2010.) Given this powerful body of evidence, it is reasonable to ask why OBPM continues to be the default discourse for those in public policy in the UK and elsewhere.

Yet literature around OBPM is replete with evidence about how difficult it is to undertake effectively. Studies have demonstrated problems concerning how to define indicators for outcomes that are meaningful but that do not distort practice, how to allocate responsibility for shared outcomes, the distorting effects of performance management, and unintended consequences.

There are two key areas where practitioners and academics have found evidence of serious problems with the implementation of OBPM. This article examines how these problems raise concerns about OBPM’s underlying conceptual framework.

**The information gap: ‘outcomes’ don’t measure impacts in people’s lives**

The first issue goes to the heart of outcomes thinking. How do we know what impact an intervention has had? The literature is full of advice for generating appropriate outcome indicators and measures (for example Liu et al., 2010; Wimbush, 2011).

However, the discussion about outcome measurement largely avoids a crucial point about the nature of measuring impact. The meaning of any ‘outcome’—the actual impact on the life of the person experiencing it—is heavily dependent on the particular context of that individual or group. The real outcome of a social policy intervention—the meaning of any change to that person—can only be understood in relation to the complexity of their lives (Widdershaven and Sohl, 1999). This is illustrated through a hypothetical case study in figure 1.

There is a paradox at the heart of outcomes-thinking concerning information about outcomes. The reason that outcomes are attractive, and more effective as a measurement tool than inputs or outputs, is that they are supposed to refer to impacts in the lives of real people (Schalock, 2001). Unfortunately, it is extraordinarily difficult and resource intensive to understand impact in people’s lives: the depth of knowledge that is required to understand the meaning of an outcome within an individual’s context demands the use of both intensive quantitative and qualitative research.
methodology with subjects over prolonged periods (Schalock, 2001).

As a consequence of this complexity and expense, those measuring outcomes turn to proxy indicators or ‘tracer conditions’ (Smith et al., 1997)—measures which focus on information that substitutes for the outcome itself, and are cheaper and simpler to collect. This makes collecting outcome information efficient and practical, and gives simple, comparable data over time (Schalock 2001; Liu et al., 2010). However, the price that is paid for this simplicity is that the depth of understanding of impact within context is stripped away.

Therefore, this argument suggests a crucial underlying conceptual flaw in OBPM. Our desire for outcome information outstrips our ability to provide it. Information about outcomes can either be simple, comparable and efficient to collect, or it can be a meaningful picture of how outcomes are experienced by people. It cannot be both.

Outcomes are complex and are not ‘delivered’ by an organization or programme
The second area of concern is the way that information about outcomes is used to make judgements. Information about outcomes is used in a variety of ways to make decisions about performance and the allocation of resources: as management information to improve the quality of services, and as feedback for policy makers to judge between the effectiveness of different types of intervention. For outcome information to be valid in any of these contexts, the problem of attributability must be resolved: those using information about outcomes must be able to understand the extent to which the interventions which they deliver/have paid for caused the outcomes that have occurred (Schalock, 2001).

The fundamental problem in this area was highlighted by Mayne (2007): ‘Outcomes are by definition results over which organizations do not have complete control’. Once more, the debate within the literature on this area has treated this issue as a practical problem which can be overcome by refining the methodology. Tactics include the use of statistical tools such as regression analysis, and by reducing the impact of ‘external factors’ on evaluation analysis by simplifying the scenarios that are considered (Schalock, 2001). Similarly, the use of so-called ‘program logic’ models is recommended, by which causal chains from specific interventions to long term impacts are mapped out (Schalock and Bonham, 2003).

However, these mechanisms fail to solve key aspects of attributability problems. Statistical analysis can only address attributability issues which were known and built into the data gathering and evaluation mechanism from the beginning. Therefore, any change to the context of programme delivery which was not identified prior to programme design will make that information significantly less useful in terms of analysing the performance of that programme. The key conceptual flaw of this approach is that it is based on the idea that outcomes are the result of a linear process from problem through intervention to positive outcome. The linear nature of this thinking is a very poor way to conceptualize an individual’s journey through the wealth of experiences, opportunities, constraints and choices which create an outcome in their life—see Ryan’s story (figure 2).

What would an OBPM approach make of Ryan’s journey? Would the housing organization have continued to support him had they been on a ‘payment by results’ model? In terms of ceasing his offending behaviour, how would it analyse the contributing role of the girlfriend?

The interaction of these conceptual flaws: social policy’s uncertainty principle
We have already seen that is impossible to create information about outcomes that is both meaningful to those who experience it and efficient to collect and analyse. We can now also see that the more information we collect about an outcome, the more it becomes subject to complexity. If we try to minimize the information gap by recording detailed information about a service user’s life, and we do so for an extended period of time following that intervention to create a decent longitudinal understanding of impact, then we multiply the range of ‘external’ influences which are creating the ‘outcome’ we are observing. The longer we spend getting to know a person, the less likely we are able to attribute any ‘outcome’ to one particular cause.

Figure 1. Understanding impact.
Let us imagine that at the end of the UK government’s Work Programme (which focuses on those who are most disadvantaged and furthest from the labour market), a person who had previously had mental health problems which prevented them from seeking employment, now has a job. This is an ‘outcome’ which we can measure relatively straightforwardly. However, discovering the genuine impact of this outcome in the life of the person experiencing is much less straightforward. Whether this is a positive outcome for that person (and society) is heavily dependent on a range of complex factors in that person’s life. How has the job impacted on their mental health? (Are they getting positive benefits from feeling useful and associating with others, or is the stress of the role damaging them?) How is the job impacting on their relationship with their dependent family? (Is the time spent on the job preventing them from playing an effective parenting role? Or is the job providing a positive role model for the family?)

© 2013 THE AUTHOR
JOURNAL COMPILATION © 2013 CIPFA
How outcome information distorts organizational priorities and frontline practice

Managers operating under OBPM performance management regimes are tasked with delivering ‘outcomes’ that are almost certainly beyond their control. What happens when managers are asked to create the required outcomes data by altering the things that are within their capacity to control. Perrin identifies that the use of outcome targets for performance management causes:

...goal displacement...which leads to emphasis on the wrong activities and encourages creaming and other means of ‘making the numbers’ without improving actual outcomes. As a result, they frequently distort the direction of programs, diverting attention away from, rather than towards, what the program should be doing.

Similarly, van Thiel and Leuww (2002) draw attention to the way in which organizations using OBPM focus their efforts on those who are easiest to help and prioritize the aspects of a service which are to be measured.

If these issues are known, why have they not been addressed? Again, the problem appears to be that they have been conceived of as technical challenges in measurement or management which can be overcome. Therefore, authors such as Perrin (1998) highlight the ways in which the distorting impact of OBPM measures can be minimized and similarly, Mayne (2007) suggests that ‘the right incentives’, together with regular changes of indicator, can lessen the danger of ‘distorting behaviour’.

However, none of these suggested responses address the underlying problem. If people are asked to manage things outside of their control, then they will learn to manage the things they can control. This task can be made more complex by ‘reviewing and updating’ indicators over time (in which case, how can those changed indicators give a useful year on year comparison of the service?) but managers will simply adapt their behaviour to meet the new targets (whatever mix or balance of indicators is chosen).

In order to effectively address this problem of the ‘performance paradox’ (Thiel and Leuww, 2002), we must stop treating it as a technical issue concerning the implementation of OBPM, and instead acknowledge that it is a problem intrinsic to OBPM as a system of performance management.

In addition to distorting the way in which organizations set priorities, there is evidence about the way in which OBPM distorts the behaviour of frontline staff, to the detriment of the people they are supposed to serve. Keevers et al. (2012) looked at the ‘before and after’ of how the introduction of a results-based accountability (RBA) approach affected the practices of frontline staff within social support organizations. Following the introduction of an RBA reporting system, staff were found to spend time collecting and analysing data about young people, rather than spend time developing and maintaining the quality of relationships with young people. As a consequence:

Both the managers and the youth workers claim that the accountability and monitoring requirements of the program...[are] paradoxically making it more difficult to engage with and build relationships with homeless and at risk young people. The integration of this new system has significant impacts on the daily practice of workers, reducing the time available to create a sense of belonging and develop young people’s skills that will form the basis for sustaining their life changes (Keevers et al., 2012).

These conclusions have been mirrored by other studies which have analysed the impact of target-driven cultures on the behaviour of...
social welfare staff (Cottom, 2011). Instead of developing relationships which enable workers to identify and meet the needs of people, workers start to see the people they serve as mechanisms to generate the data they need to prove that they are being effective.

**Conclusion**

The evolution of public policy thinking towards a perspective which attempts to understand and value interventions in relation to the impact they have on people’s lives was both necessary and welcome. However, OBPM, and particularly ‘payment by results’, create a performance paradox: an outcomes based approach distorts the priorities and practices of those tasked with delivering such intervention and results in worse outcomes for the people who are supposed to benefit.

We have used ‘hard cases’ to exemplify problems with OBPM as a conceptual framework. It could be argued that any framework fails if you provide it with difficult enough problems. However, this is an unsatisfactory response, for two reasons. First, social policy interventions are intrinsically complex and ‘hard’, as their currency is the complexity of lives as they are lived. Social policy frameworks must be able to deal with ‘hard’ cases. Second, it is not just the hard cases where OBPM fails. Every time a person’s life is reduced to easy-to-capture data, it is a failure of OBPM, because some aspect of that representation will be inaccurate. Every time an organization is paid/not paid for delivering an ‘outcome’, a multitude of complex factors will have been ignored. Every time a manager makes a decision to do what makes the figures look good, rather than what is actually needed, it is a failure of OBPM. Every time a teacher teaches to the test or selects only the bright pupils to sit an exam, it is a failure. Every time a worker captures data rather than listening to a client’s story, OBPM has failed.

We need to stop treating these as practical problems to be resolved, but as evidence of conceptual flaws in OBPM. It is time to develop a new theory. Our challenge is how to maintain the focus on beneficiaries which OBPM was supposed to deliver, without replicating its distorting effects on practice. Promising avenues in this search include aspects of systems theory (see, for example, Williams, 2006) and approaches which are grounded in theory about how people make choices in complex environments (see, for example, Archer, 2003). Now is the time to heed W. Edward Deming’s words (quoted in Perrin, 1998):

> Focus on outcomes (management by numbers, MBO, work standards, set specifications, zero defects, appraisal of performance) must be abolished, leadership put in its place.

**References**


