



IJPPM
54,3

212

Received November 2004
Accepted November 2004

PROFESSIONAL PRACTICE

Effective review meetings: the counter-intuitive key to successful performance measurement

Alan Meekings

Landmark Consulting, London, UK

Abstract

Purpose – To present clear pointers to best practice in the implementation and use of performance measurement, highlighting the significance of effective review meetings, connected across functions and between levels in organisations.

Design/methodology/approach – A summary of practical experience in the implementation and use of performance measurement, supported by three complementary case studies, each illustrating ways in which it can be both a catalyst and impetus for results otherwise unachievable. In particular, this paper explains why the greatest benefits lie not in specific measures themselves, nor even in the process used to develop them, but in how performance measurement is actually implemented and used in practice. It argues that the key to getting the most out of performance measurement is to grab the review process and make it work properly.

Findings – Insights into how to maximise the value of performance measurement by making the review process work properly, such that strategic developments and continual improvement are effectively “pulled-through” in a coherent and progressive fashion and the measures themselves become essentially “self-correcting”.

Originality/value – This paper brings performance measurement to life, by concentrating on how it is best implemented and used in practice through focusing on what it takes to make the review process work properly from top to bottom and side to side across organisations. It also offers some interesting hypotheses for further academic work in this field.

Keywords Performance measurement (quality), Case studies

Paper type General review

The context

For over a decade, since the publication of Thomas H. Johnson’s book *Relevance Regained* (Johnson, 1992) and Kaplan and Norton’s first “balanced scorecard” article (Kaplan and Norton, 1992), the focus of attention among academics and practitioners has centred on the determination of appropriate performance measures. As Professor Johnson said (Johnson, 1992, p. 9), “[M]anaging by remote control with accounting-based information perpetuates practices that contradict improvement strategies associated with competitiveness”. Hence, the hunt has been on for “better measures”.

During the intervening years, further frameworks, such as the “performance prism” (Neely *et al.*, 2003), have been developed to inform the design of measurement systems,



and the original balanced scorecard framework itself has undergone several generations of enhancement (Cobbold and Lawrie, 2002).

As more and more organisations have implemented new performance measurement frameworks, so the benefits of going through the implementation process – in terms of greater executive alignment, improved communication of strategy, and so on – have come to be recognised as valuable in their own right. Indeed, it was David Norton who once observed that “[D]eveloping a balanced set of measures is important, but the process an executive team goes through in developing these measures is much more important still” (Norton, 1995).

However, the *greatest benefits* from the application of performance measurement lie not in measures themselves, nor even in the process used to determine them, but in how they are actually implemented and used in practice. Counter-intuitively, the key to success lies in doing everything necessary to make the review process work properly, and connecting it effectively across functions and between levels in the organisation.

This paper presents a number of clear pointers to best practice, derived from years of practical experience, and also offers some interesting hypotheses for future academic research.

The challenge

Unfortunately, most people’s experience of performance measurement in organisations is not encouraging. How many demonstrably successful organisations would claim that performance measurement has been the key to their success? Few, if any. Indeed, the way performance measurement is used, or not used, is typically more of a millstone than an enabler for strategic development or fundamental improvement.

Talk to managers and front-line staff in most organisations, even those who believe they have implemented leading-edge performance measurement frameworks, and you are likely to encounter anything from apathy to disdain on the topic. Several common concerns emerge:

- long, unproductive management meetings are the norm, with little or no purposeful action resulting;
- strategic measures are disconnected from operational issues and seem irrelevant in the context of day-to-day pressures to “deliver the budget”;
- organisations are drowning in data, yet derive little, if any, insight from their measures;
- arbitrary numerical targets distort collective behaviour and seriously damage the way staff and managers view their work and their organisations;
- sub-optimisation occurs within functional departments to the detriment of overall organisational performance; and
- fear of personal exposure conceals the facts, and works against creating a genuinely open learning culture.

This is not a pretty picture. As professionals with an interest in this field, though, we need to be honest about current reality.

Fortunately, the answer is surprisingly simple and elegant. Critically, it is grounded not in specific measures *per se*, nor in how they are determined, but in how performance measurement is actually implemented and used in practice.

The answer

The answer to getting the most out of performance measurement is to *grab the review process and make it work properly*. The review process is “where the rubber meets the road”. Without effective review, nothing happens; and without connecting the review process effectively between levels, issues cannot be addressed appropriately or systemically.

The beauty of focusing on the review process is that everything that needs to be done to make the review process work properly is effectively “pulled through” in a coherent, progressive fashion, rather than having to be “pushed at” the organisation as a series of disconnected initiatives.

Moreover, provided the review process is made to work properly, experience shows that measures themselves become essentially “self-correcting”, in the sense that any initial, “provisional” measures that deliver no useful insights will be dropped as other more appropriate measures emerge.

To make the review process work properly, organisations need to:

- structure the review process properly, defining who needs to come together to review “what, when, why, where and how” – in particular, the frequency of measurement and review is critical to success;
- connect the review process effectively between all levels in the organisation;
- ensure that review meetings deliver value in their own right – this in turn depends on optimising the value of measurement, and presenting data visually in a way that aids intelligent interpretation, particularly by making variation visible and by highlighting ‘signals from noise’ in large volumes of data; and
- tackle implementation in a manner that paves the way for sustainable use.

These recommendations will be illustrated by three disparate yet high-profile case studies, each demonstrating, in its own way, how performance measurement can be both a catalyst and an impetus for results that are otherwise unachievable.

Case studies

The three case studies are:

- (1) the improvement of public services in New York City under Mayor Giuliani’s eight-year tenure from 1 January 1994 to 31 December 2001;
- (2) the turnaround of a manufacturing plant in Scotland facing imminent closure in the early 1990s, followed by the subsequent transformation of performance at Rolls-Royce Motor Cars during the mid-1990s, based on similar principles; and
- (3) the dramatic, and publicly recognized, improvement of the Automobile Association’s Roadside Services – “the fourth emergency service” in the UK – over recent years.

New York city

Rudolph W. Giuliani’s recent book (Giuliani, 2002) describes how, as Mayor, he led New York from what had been termed “ungovernable” status to achieve some of the most impressive improvements in public service delivery seen anywhere in the world. Chapter 4 of his book, in particular, explains how the City delivered a string of unprecedented improvements, such as:

- cutting its crime rate by 57 per cent in less than five years;
- almost eliminating stabbings and slashings in its prisons, specifically reducing the average number of incidents from 139 a month in 1996 (the worst incidence of inmate-on-inmate violence for any US city) to less than one a month in 2001; and
- dramatically lowering the number of cases per Child Support Specialist, from 28 in 1996 to less than half that five years later, while simultaneously increasing the rate of successful child adoptions by 66 per cent over the same period.

These improvements were typically accomplished with no increases in funding or changes in legislation. So how were they achieved?

Giuliani's own answer to this question can be summarised thus:

- start by measuring issues that matter to citizens, customers or consumers, with data ideally captured daily;
- rigorously review, ideally weekly; and
- hold individuals accountable for outcomes.

Note that no arbitrary numerical targets were set and yet huge, previously inconceivable, gains in performance were made.

Measuring what matters to citizens, customers or consumers. Giuliani wanted results, so he insisted that his various service departments should measure outcomes in end-user terms. The Police Department were the first to pilot what became known as the "Compstat" system. Formerly, the Department had focused on statistics such as the number of arrests or the time to respond to emergency calls. Their new measures were all related to the fundamental purpose of the Department, namely "public safety and reducing crime". Financial and HR measures were eventually added, but much later, and the number of measures used was not artificially constrained. The Department of Corrections, for instance, reviewed 592 indicators weekly[1]. Indeed, one of the keys to success in New York was capturing relevant measures daily and then reviewing them weekly, in line with what could be termed "the heartbeat of performance".

Rigorous review. In the Police Department, Giuliani introduced the concept of a "floor plan", specifying precisely who had to attend each weekly review meeting. The review process itself was rigorous and relentless, yet sparked the insights that subsequently informed success. As he says in his book, "[E]ven after eight years, I remain electrified by how effective those Compstat meetings could be".

Accountability. The senior managers from all eight Police Precincts attended each weekly review meeting. On the one hand, "commanders had objective proof of their good performance" and were able to learn from what seemed to worked elsewhere. On the other hand, there was no hiding place, nor escape, from individual accountability for outcomes. However, by emphasising results rather than methods, managers were given considerable latitude in experimenting with their own means of improvement.

A manufacturing plant in Scotland, and Rolls-Royce Motor Cars in Crewe

A manufacturing plant in Scotland was facing imminent closure. The challenge was to achieve dramatic improvements in quality, delivery and cost within a three-month period, or else the factory would be shut down.

The single most important change made was the implementation of a completely restructured review meeting framework, linked from top to bottom and spanning all functions in the factory.

In order to inform the reorganisation of the factory floor, the flow of work was mapped horizontally – “from left to right” – using conventional process flow charting techniques. In contrast, to inform the new review meeting framework, the flow of information was mapped vertically – “from bottom to top” – both as it was (the “as is” situation) and as it needed to be in future (the “to be” situation).

Essentially, each level of review looked at three types of measures, together contributing to a genuinely systemic picture:

- (1) “tactical” measures of operational performance;
- (2) “strategic” measures concerning development, innovation and continual improvement at each particular level; and
- (3) “audit” measures of interest to the level above.

Each level tracked only as many measures as necessary to present a systemic picture of performance, and no more.

At the outset, there were seven hierarchical levels of management in the factory, whereas the “to be” information flow starkly illustrated that only three levels of review were necessary:

- (1) daily;
- (2) weekly; and
- (3) monthly.

Consequently, the review meeting architecture specified only three levels of review, detailing who needed to come together to review “what, where, when, why and how”, regardless of the number of levels of management currently in place. There was simply no need to have seven levels of review just because there were seven levels of management.

This new review meeting architecture was then systematically implemented. The impact was dramatic and sustained. Factory performance, in terms of quality, delivery and cost, was transformed within three months, and factory closure was avoided.

When the opportunity arose, shortly after, to enable an equally dramatic turnaround at Rolls-Royce Motor Cars in Crewe, a similar approach was adopted. The factory was reorganized into zones and teams, and an appropriate structure of daily, weekly and monthly review meetings was put in place.

The time taken to build a Rolls-Royce or Bentley motor car was subsequently cut from around 80 days to under a month. This enabled production to be shifted from a basis of “make to stock” to “make to order”. Quality, delivery, cost and profitability improved dramatically, so much so that Volkswagen and BMW subsequently paid a high price to acquire the business.

Continual improvement became ingrained, underpinned by the new framework of measurement and review. Cars are now currently built in the same factory in less than three days.

Two additional factors are worth highlighting:

- (1) The person at the “top” (at any level in the organisation) must lead and role-model the review process. Without the person in charge showing that effective review matters, little benefit will result. At Rolls-Royce Motor Cars,

- each functional Director personally led the review of departmental performance, as did the Managing Director for collective performance.
- (2) The frequency of review is highly significant. It typically takes around six periods of review to reach what might be termed the “ah ha” point – the point at which people can connect the information they receive with the decisions they take, the actions they implement and the resulting outcomes (see Figure 1). So, with weekly review, it usually takes around six weeks for the necessary connection to be made. Interestingly, if data merits daily review, the “ah ha” point can be reached in less than six days. In contrast, with monthly review, it may take much longer than six months, and, with annual review, it may never happen – simply because the longer the gap between insight, decision, action and outcome, the more difficult it is to make the necessary connection between insight and outcome (see Figure 2).

Automobile Association (AA) Roadside Services

The British public has always held the AA in fond regard. This dates back to its founding, in 1905, as a “body intended to help motorists avoid police speed traps”.

However, as the market for roadside services in the UK grew, so other operators began to challenge the AA in terms of price and customer perceptions. Even *Which?*, the well-known UK consumer magazine, stopped recommending the AA as its “best buy” for breakdown services. The situation today, though, looks very different. Indeed, the most recent *Which?* survey of breakdown services (*Which?*, 2002) recommends readers to switch back to the AA, and the J.D. Power survey of customer satisfaction is equally complimentary (J.D. Power and Associates, 2002).

Reaching the ‘Ah-Ha’ Point

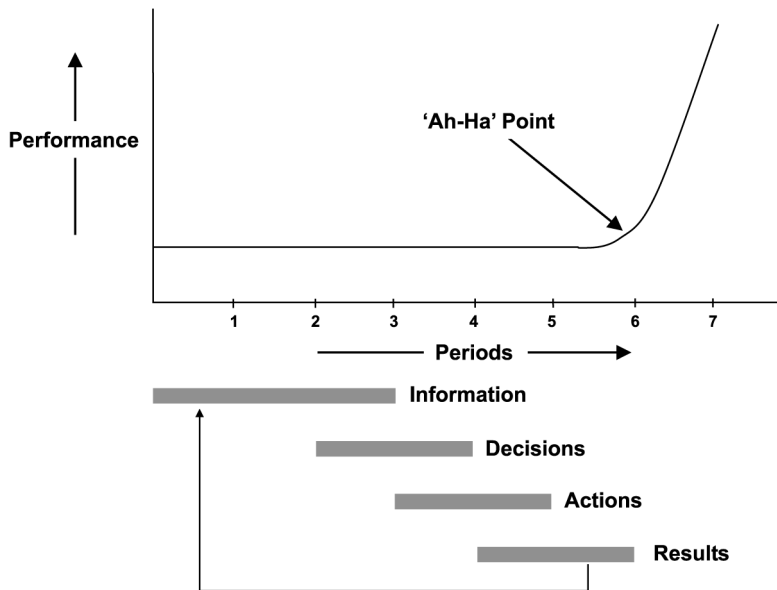


Figure 1.
Reaching the “ah ha” point

The Significance of Review Frequency

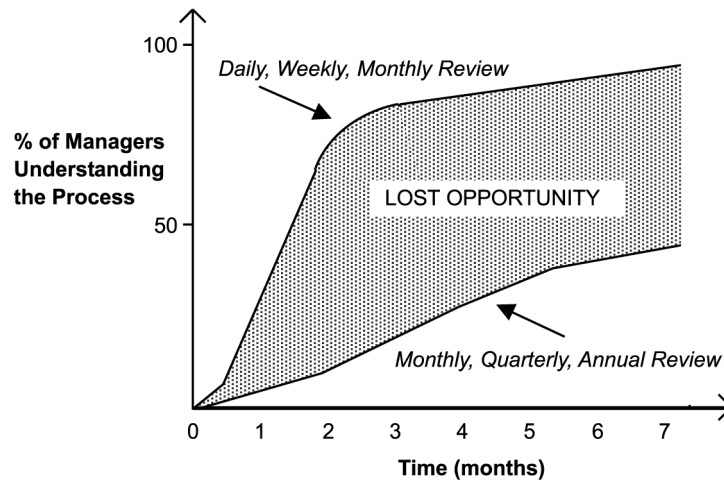


Figure 2.
The significance of review frequency

Changing the measurement and review framework at AA Roadside Services played a major part in its transformation. The key changes (Wood *et al.*, 2001) were:

- using statistical process control (SPC) charts for management information, both at activity level and in summary format (as “performance dashboards”), thereby making variation visible and enabling everyone to filter “signals from noise” in their data;
- getting the review process to work effectively at each level in the organisation (in this case at the front line, within the regions and at the centre);
- connecting the review process between differing levels so as to create a beneficial “updraught of management attention”, drive the process of strategic development, innovation and continual improvement, facilitate the process of best practice transfer, both locally and across the whole operation, and allow issues to be addressed at the most appropriate level from a systemic perspective; and
- tackling implementation in a way that allows performance measurement and review to take root and prove its value – despite the many recognisable reasons why people tend to resist making performance explicit and widely visible: for example, the new measurement and review process at the AA was initially piloted with a high-performing team best able to grasp the concept and run with it.

Key themes

Six key themes emerge from these disparate case studies. They shed an interesting light on how to maximise the value of performance measurement in organisations, even though many of them run counter to current “received wisdom”. They would serve as useful hypotheses for future academic research:

- (1) *Optimise the value of measurement* – to optimise the usefulness of measurement, it is essential to include measures that matter to citizens,

customers or consumers, the coverage of measures needs to be sufficiently comprehensive to provide a genuinely systemic picture of performance, and data needs to be captured sufficiently frequently to track the “heartbeat of performance”.

- (2) *Present data visually to aid intelligent interpretation* – to derive the maximum value from data, all process-based data needs to be presented in time-series format, ideally showing the limits of predictable variation, and large volumes of data need to be presented in a “performance dashboard” format to highlight the very few insights economically worthy of attention.
- (3) *Structure the review process properly* – to structure the review process properly, there needs to be a clear definition of who needs to come together to review “what, where, when, why and how”, and the frequency of review is critical. Typically, the faster the cycle of review, the faster the rate of progress.
- (4) *Ensure that review meetings deliver value* – review meetings will deliver value provided that the measures themselves are useful (see point 1, above), the data is presented in a visual format that aids intelligent interpretation (see point 2, above), the person accountable makes the review process matter, no arbitrary numerical targets are set (goals are fine, as long as the culture promotes root-cause analysis, problem-solving and innovation, whereas arbitrary numerical targets are not), and people need to be held accountable for outcomes.
- (5) *Connect the review process between levels* – connecting the review process between levels means promoting problem-solving at the local level, putting in place a mechanism for best practice transfer (particularly in large organisations), creating a mechanism for ideas (and outline business cases, where appropriate) to be submitted upwards for incorporation into properly planned and resourced improvement initiatives, sponsored at top level, with feedback flowing to individuals who suggested them, and allowing issues to be addressed at the most appropriate level of review from a systemic perspective.
- (6) *Tackle implementation progressively* – implementation will succeed and take root if the approach adopted is more like doing a jigsaw (i.e. tackling the easiest pieces first) than moving forward on a broad front, the initial measures adopted are clearly positioned as “provisional” (i.e. open to revision in the light of subsequent experience), and the person at the top of the organisation pays serious attention to grabbing the review process and ensuring it works properly – up, down and across the organisation.

Summary

At first sight, these points may appear to be obvious. Yet optimising the review process is seemingly the last thing that occurs to most organisations. Typically, most senior managers are either exhausted by the process of defining and implementing their new measures or they interpret any offer of help with their review process as some sort of implied criticism of their managerial skill.

Nor do most management consultancies help in this respect. It is clearly much easier to persuade a CEO that the measures used in his or her organisation are less than fully balanced (let alone systemic), and hence sell them a project to implement a new measurement framework than it is to convince them that they need help in changing

the way they manage their organisation. How many CEOs do you know who would confess openly that their management methods are in need of serious improvement?

Those who wish to coach organisations in getting the most out of performance measurement need to have experienced for themselves “the art of the possible”. They also need to understand how to take organisations through the steps necessary to reach this point. There is simply no substitute for practical experience in grabbing the review process and knowing what to do – in priority order – to make it work properly. The prize is far greater than most people would imagine.

Perhaps the quickest way to propagate the recommendations outlined in this paper would be for the academic community to test them as hypotheses and then popularise their findings, in much the same way as Kaplan and Norton did with the “balanced scorecard” over a decade ago.

Note

1. Most current academic literature on performance measurement recommends reviewing only a limited number of strategically-significant indicators at senior level – certainly not as many as 592 for a single function – and yet New York’s achievements were truly “strategic” by any yardstick.

References

- Cobbold, I. and Lawrie, G. (2002), “Development of the balanced scorecard as a strategic management tool”, in Neely, A., Walters, A. and Austin, R. (Eds), *Performance Measurement and Management: Research and Action*, Centre for Business Performance, Cranfield, pp. 125-32.
- Giuliani, R.W. (2002), *Leadership*, Little Brown & Co., Boston, MA, pp. 69-97.
- J.D. Power and Associates (2002), “UK Roadside Assistance Study”, June, available at: www.jdpower.com
- Johnson, H.T. (1992), *Relevance Regained: From Top-Down Control to Bottom-Up Empowerment*, Free Press, New York, NY.
- Kaplan, R.S. and Norton, D.P. (1992), “The balanced scorecard: measures that drive performance”, *Harvard Business Review*, Vol. 70, January/February, pp. 71-9.
- Neely, A., Adams, C. and Kennerley, M. (2003), *The Performance Prism: The Scorecard for Measuring and Managing Stakeholder Relationships*, Financial Times-Prentice-Hall, London.
- Norton, D. (1995), Video conference, 9 March.
- Which?* (2002), “Breakdown services”, *Which?*, November.
- Wood, B., Williams, R., Anker, D. and Gardener, N. (2001), “Performance management and improvement at the Automobile Association”, *Supply Chain Practice*, Vol. 3 No. 1, pp. 30-42.

Further reading

- Greif, M. (1991), *The Visual Factory: Building Participation through Shared Information*, Productivity Press, Cambridge, MA.