

## **Performance measurement systems and performance measures: the good, the bad and the ugly**

*Paper delivered to the Australasian Evaluation Society Conference, Wellington New Zealand, Friday September 3 2010*

*Graham Smith  
Numerical Advantage*

This paper is based on the author's more than 10 year experience in developing performance measures in a variety of settings, mostly Australian Government organisations. One of the main things learnt in this time is the need to think differently about the performance measure system and the performance measures themselves.

Performance measures are key elements that link evaluators, commissioners and users of evaluation. Measures are, or should be, a reflection of reality, and their quality (the extent to which that reflection is clear) is key.

There are various acronyms or guides for assessing performance measures. One of the more popular is the SMART acronym: that measures should be specific, measurable, achievable, relevant and time-based.

But that only talks about the indicators themselves. It says nothing about the organisational context in which they will be deployed, the extent to which they will be supported by staff, management and stakeholders, the 'infrastructure' – e.g. the extent of reasonable sources of good quality data and the history on dealing with things quantitative, and even the ethics of the organisation – which can influence, for example the degree of trust and the likelihood that indicators will be gamed.

So in order to understand the way in which performance measures work, we need to consider the system in which they reside.

### **Need for performance measures**

First, let us consider the main reason for defining and collecting performance measures: performance improvement, accountability and decision-making.

*Performance improvement* can be achieved through performance measures that are then compared with a standard or benchmark in order to reveal opportunities for improvement.

*Accountability* can be achieved through comparing the performance measures achieved with those that were promised – typically in a budget document. It can also be considered as a more general obligation to meet the needs of internal and external stakeholders.

*Decision-making* can be assisted through comparing performance measures achieved by competing programs that are directed to the same objective. Decision-making can also include rewarding individuals for performance, or for adjusting resourcing based on performance. This can go both ways: poor measured efficiency may lead to funds

being withdrawn; poor outcomes for the community may, in some circumstances lead to the need for additional resources to be invested. Either way, good performance measures are needed to support the decision.

By extension, performance measures can assist with prevention of waste. Together with improved accountability, they also contribute to improved government and hence confidence in social structures. So they are potentially powerful instruments; and as such they can be used for evil as well as good..

### **Hazards of performance measures**

The case against performance measures includes that they can be subject to gaming, have problems of immeasurability, can be costly and time-consuming to introduce, and cause embarrassment and loss of power. These are now discussed briefly.

#### *Gaming*

Gaming refers to the practice whereby activity is directed towards achievement of the performance measure, in such a way that the overall objective is not likely to be achieved. This especially arises where the performance measure is at the process or output level rather than an outcome itself. Examples include:

- 'teaching to the test' Exams and other means of student assessment are imperfect performance measures of the knowledge a student has gained through attending a class. The measure is further damaged if the teacher focuses purely on those precise elements that are testable.
- short-term measures of success. For example, an organisation that rewards its senior officers based on this year's profit can encourage the taking of short-term views and even distortion of the profit reported.
- poorly defined measures that can be gamed. For example, a performance measure that required 90% of tasks to be completed in 30 days was gamed through the area responsible expediting tasks that had reached 25 days, and abandoning those that were likely to miss the target. A task that took 31 days was, as far as the performance measure was concerned, as bad as one that took 131 days.

Campbell's Law states that: "The more any quantitative social indicator is used for social decision making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor. However, this is not to say that quantitative indicators should not be used, but that they should be used with caution.

This paper emphasises the systems approach, whereby a single measure may well be abused in the wrong circumstances, but that several measures working together can give a more reliable view. For the example above about 30-day tasks, two other performance measures assisted in solving the distortion. First, the level of complaints rose (always a useful measure to take into account). Subsequently, an additional performance measure, namely the age of the oldest outstanding task, was introduced. The solution to a corrupted performance measure may well be another performance measure!

### *Immeasurability*

It is claimed that some things are immeasurable. That may be so, but they are given 'weight', and need to be balanced against other criteria, thus acquiring a *de facto* measure. For example, decisions made on road safety or health expenditure imply a value given to a human life.

A second source of immeasurability is the long-term nature of some outcomes, which means that they cannot be measured directly. For example, improvements to the natural environment, or outcomes in educating a child for life, will not be known for many years. But decisions still need to be made; implying the need for a proxy measure such as an output (ideally with good science linking it to the outcome).

The author's view is that the assessment of an outcome should be as clear as possible, and this is often (but not always) achieved by enumeration. As Galileo said, 'Measure what is measurable, and make measurable what is not so.'

### Cost and Time

Like any other management activity, the creation and use of a system of performance measures needs to earn its place by being cost-effective. One view is that focus on performance measures can displace activity from more considered judgements and evaluations. However, measurement and evaluation are not exclusive, and enumeration can assist in other forms of assessment.

One aspect of cost is to have effective reuse of performance data for both local decision-making and for integration as whole-of-organisation performance measures. Another aspect is to not have an excessive number of measures.

### *Embarrassment*

This is the converse of accountability. Some persons who have not previously been held accountable for results may only see in the production and use of performance measures the potential for embarrassment.

### *Loss of Power*

Without a system of performance measures, those who are well-connected are best placed to have information on the performance of the organisation. As information is often power, the introduction of a system of performance measures may be seen to impinge on their power.

## **Dimensions of performance measurement systems**

The nature of performance measurement systems can be described using a number of dimensions or characteristics. These characteristics, which will be described in more detail in the following paragraphs, are:

- Ownership of the performance measurement system
- The purpose and use of the performance measurement system
- How the system is designed, coordinated and resourced

- Whether targets for the measures are specified, and if so, of what type
- Whether and how the performance measures are linked to strategic planning
- Whether there is independent review of the performance measurement system.

### *Ownership*

One of main issues is ownership of the performance measurement system. An owner is someone who can make the decision to cancel or delete the performance measurement system. Some of the possible variations here are:

- Nationally mandated performance measures – e.g. Productivity Commission
- Jurisdictional performance measures – e.g. performance measures set through the budget process, or legislated performance targets.
- Organisationally mandated performance measures, e.g. those established within organisations, for example in strategic and business plans.
- Locally mandated measures, for example benchmarking efforts by specific groups.

Performance measures that are established at the highest level are not necessarily the 'best'. It depends on the purpose. For example, the PC measures are useful for overall performance comparisons at a jurisdictional level. They can be at a little bit of a gross level, but are sufficient to at least provide an indication of directions to pursue to improve performance. At the other end of the scale, benchmarking exercises are less public but more targeted to specific improvement.

### *Purpose and Use*

Purposes for performance measures: accountability, performance improvement and decision-making were discussed above in the section on 'need for performance measures'.

These purposes for measures can also be extended to be purposes for performance measurement systems. However, there are also some other purposes for the overall system of performance measurement that can be more hidden and perhaps have negative connotations. These include justification, avoidance and reform, discussed below.

*Justification.* Some performance measurement systems may allow or encourage areas within the organisation to propose measures that justify the existence of that area, rather than being reflective of overall organisational performance.

*Avoidance.* In some organisations, there may not be a commitment to performance measurement, but it is nevertheless imposed from a higher authority (see 'ownership' above). In these circumstances, the organisation may have the aim of avoiding accountability by doing as little as possible with performance measures – selecting easily measured, easily achieved measures that may not assist overall assessment.

*Reform.* On some occasions, for example, when there is a reorganisation or a new leader, performance measures may be used as an instrument to drive reform. Top-down designation of what the measures and their targets are can be one of the means to achieve a change in direction.

### *Design, coordination and resources*

A performance measurement system needs effort to design and coordinate. The main issues are the selection of the performance measures and the processes to collect, review and present the results. These functions may be carried out primarily by line areas within the department (bottom-up) by senior management (top-down) or through a specific coordination unit.

A line area may choose performance measures that are too easily achieved, and hence not appropriate if the overall objective is accountability. On the other hand, measures selected by the line area are more likely to be well-targeted if the objective is a line-area driven improvement or benchmarking process. Conversely, a top-down approach (or at least strong top-down review) may be more appropriate if the objective is accountability.

Collection of data for performance measures may be explicitly collected by the line area, in which case (especially when collected for accountability purposes) the data will need independent review or audit. In some cases, the performance measures can be extracted directly from pre-existing management information systems which, depending on the controls on those systems, can produce reliable results.

Design and coordination of performance measures by a central advisory group can be effective, especially when there are multiple roles for performance measures to play. Such a group can help to assure senior management on the validity of measures and the data and assist line areas in data collection and analysis of local improvement purposes. Such an advisory group, however, comes with a cost.

The effective functioning of a performance measurement system also requires training and advice on the collection of data and the reporting of performance measures. This is often more effective if provided centrally, by a group such as the advisory group described above. Other sources of information can be the finance function, but in some organisations there is little advice provided at all.

### *Targets*

The first issue is whether the performance measurement system provides for formal targets. In some cases, the emphasis is on simply measuring the performance that is achieved, without a target as such. This is often appropriate when the measure is such that there is no natural limit - more is always better. In these cases, however, there tend to be informal targets - to achieve a better result than comparable organisations (benchmarking) or to achieve a better target than the previous period (trends).

Where there are formal targets, they can be negotiated or imposed through mechanisms such as the strategic plan or the budget. The targets themselves may be stretching or conservative; and be either soft targets (good to achieve if possible) or mandatory targets, such as those set by legislation.

### *Linkage to strategic planning*

The main variables here are whether performance measures and their targets are based on strategic plans, whether the plans are based on the measures and targets, or whether there is no connection at all.

In the public sector, the strategic plan often comes first. This is then a source document, which is used to derive organisational performance measures and possibly targets. Some organisations also produce business or operational plans for lower levels of the organisation. These have to be consistent with the strategic plan, and are in turn used to derive performance measures and targets.

On some occasions, especially in the private sector, the measures and targets come first. For example, an organisation may decide to aim for a particular market share or level of sales. A plan to achieve this target is then derived. In practice, of course, there is a degree of complementarity between targets and plans such that it is difficult to determine which came first.

A less effective arrangement occurs when the plan and the measures are not connected. When this occurs, there tend to be multiple directions for the organisation, leading to lack of clarity of purpose. There is also excessive collection of information to support both the performance measures and reports against strategic plan progress.

#### *Independent review*

A major problem for performance measurement systems is gaming; and another problem is obfuscation or straight lying. One response to both of these is independent review on the correctness of a performance measure. The nature of this review can include audits of departmental performance measures in audits of financial statements by Auditors-General (as in the ACT, WA and Victoria). Finance Departments usually review performance measures that are issued with budget documents; these reviews can be of greater or lesser vigilance.

#### **Conclusion – performance measurement system types**

This paper has discussed a number of dimensions that define performance measurement systems. These dimensions or characteristics are not fully independent, but tend to cluster together to form self-consistent types of performance measurement system, some examples of which are in the table below.

|                               | <b>Imposed<br/>Accountability</b> | <b>Local<br/>benchmarking</b> | <b>Internal reform</b>    | <b>Minimal</b> |
|-------------------------------|-----------------------------------|-------------------------------|---------------------------|----------------|
| <b>Owners:</b>                | External                          | Local                         | Organisation              | Organisation   |
| <b>Purpose:</b>               | Accountability                    | Improvement                   | Improvement               | Justification  |
| <b>Review:</b>                | External                          | none                          | Organisation              | none           |
| <b>Design:</b>                | Top-down                          | Bottom-up                     | Top-down                  | Bottom-up      |
| <b>Targets:</b>               | Negotiated                        | Emergent                      | Stretching                | Conservative   |
| <b>Links to<br/>planning:</b> | none?                             | Plans based on<br>targets     | Targets based<br>on plans | None           |

For example, one performance measurement system, which is labelled 'imposed accountability' above, describes a system owned and controlled outside of the organisation being measured and deployed for the purposes of accountability. In this case, there tends to be external review (to ensure the validity of the accountability) and for the design of the performance measures to be imposed from the top. Targets

will be set, but may be negotiated between the parties to the accountability relationship. With the focus on achieving specific results, there may not be a strong linkage with strategic planning.

Another purpose can be local benchmarking, in which case the performance measurement system is owned locally and directed towards improvement. Targets emerge from the benchmarking exercise. Sometimes there is little reason at all for a performance measurement system other than a perception that there should be one, or that there is no motivation to remove an existing system. In this case, labelled 'Minimal' in the table above, the purpose is sometimes reduced to justifying activity, targets tend to be conservative (not stretching), and there is little review of the system or links to planning.

### **Problems in performance measurement systems**

Each of the systems above is reasonably self-consistent, and so as a system is stable and may even meet the owner's objectives. Even the 'minimal' model listed above, though in many ways unsatisfactory, may meet the owner's objectives of saving money and avoiding scrutiny.

This paper contends that problems are more likely to arise if the characteristics of the performance measurement system are inconsistent. For example, if the purpose is accountability, but measures are designed on a bottom-up basis; or if measures are used for decision-making (and these decisions can benefit organisational staff) but the measures are not subject to independent review. Other problems can arise if there is no purpose at all defined for the performance measurement system, in which case it is unlikely to achieve anything. Similarly, although in different circumstances measures come before plans and vice-versa, if there is no connection between the two, there is likely to be at least inefficiency and possibly confusion.

### **Conclusion**

This paper has claimed that performance measures can contribute significantly to the accountability and efficiency of an organisation. Its major new perception is that it is important to perceive performance measures in a systems context, with different types of system being appropriate to different organisational purposes. Finally, performance measurement systems that are self-consistent are more likely to be successful in meeting the organisation's objectives.