

The politics of performance indicators

Denise Lievesley

UNESCO Institute for Statistics

Institut de statistique de l'UNESCO

Montreal, April 2005

yes minister

The case of the hospital with 500 administrators and no doctors, nurses or patients

Mrs. Rogers: *"Minister, it is one of the best run hospitals in the country. It is up for the Florence Nightingale Award."*

Jim Hacker: *"And what is that?"*

Mrs. Rogers: *"It is won by the most hygienic hospital in the area."*

Performance indicators: good, bad, and ugly

- Performance monitoring done well is broadly productive for those concerned. Done badly, it can be very costly and not merely ineffective but harmful and indeed destructive.

*Journal Royal Statistical. Soc. A (2005)
168, Part 1, pp. 1–27*

Performance monitoring raises the profile of data but risks bringing official data into disrepute

- *A performance indicator defines the measurement of a piece of important and useful information about the performance of a programme expressed as a percentage, index, rate or other comparison which is monitored at regular intervals and is compared to one or more criterion.*

concerns

- Performance indicators have reached **plague proportions** in NSW. They are overrated by the NSW Premier and Government. They are misunderstood and misused by Government departments. They are undervalued by the welfare sector. They are over-promoted by consultants. They are also one small important element in the development of quality services.

- The trends within NSW in the past few years can be crystallised by the State Government's slogan "Putting People First by Managing Better". With the slogan has come management by objectives, demands for higher levels of accountability and the need to sell to the electorate the fact that the state is managed better and is performing better. **The need to sell the idea that performance is improving has overshadowed the need to actually improve performance.**

Uses of indicators – both as carrots and sticks

- *“Those councils that get the highest star ratings will get significant freedoms: ... The better you do, the more you get” **BUT***
- *“Where councils are persistently in trouble and failing to deliver... central government can not stand idly by”*
- *“Next year we will be challenging the issue of councils being happy to be average”*

To legitimise central intervention

- *“We will not tolerate poor performance or failing councils. They let down the people they are elected to serve. They tarnish the reputation of the rest of local government.” Nick Raynsford 2002*
- The response from ‘Thames Marsh’
“What makes you think you can come swanning down here from Whitehall telling us how to run our borough”

- **Performance Indicators: A Summary**
- **D. Fisher, K. Rubenson, Kathryn Rockwell
Garnet Grosjean, Janet Atkinson-Grosjean
*Centre for Policy Studies in Higher Education and
Training (CHET)*
University of British Columbia**
- While there is no single, agreed-upon definition of PIs, a performance indicator is an authoritative measure usually in quantitative form of an attribute of the activity of a higher education institution. The measure may be ordinal or cardinal, absolute or comparative. The measure includes both the mechanical applications of formulae (where the latter are imbued with value or interpretative judgements) and such informal and subjective procedures as peer evaluation or reputation rankings.

The policy context in higher education

- Changes in the public management paradigm over the last two decades.
- Public sector reforms induced fundamental changes, not only in policies and practices, but also in the culture underlying public administration
- This new culture - market-like principles of cost-recovery, competitiveness, and entrepreneurship in the provision of public services
- Criteria of economy and efficiency became dominant. Demands for efficiency were supported by "broad accusations of waste, inefficiency, excessive staffing, unreasonable compensations, freeloading, and so forth"
- Bureaucrats installed 'rational' corporate management techniques like accounting, auditing, accountability, and performance models
- to normalize and entrench private sector principles

Indicators in development

- Education– *Jomtein 1990*
- Children – *New York 1990*
- Environment and development – *Rio 1992*
- Population and development – *Cairo 1994*
- Social development – *Copenhagen 1995*
- Women – *Beijing 1995*
- Education – *Dakar 2000*
- Millennium summit – *New York 2000*
- Sustainable development – *Johannesburg 2002*
- Information Society – *Geneva 2003*

The indicator business

Some Indicator initiatives

- United Nations Development Assistance Framework
- Minimum National Data Set
- UN Economic and Social Council
- Friends of the Chair of the UN Statistical Commission

Trying to manage the tension between the need for data and the response burden on countries

How important are indicators?

- High political significance
- Raised importance of measurement
- Valuable for accountability, monitoring and advocacy

but

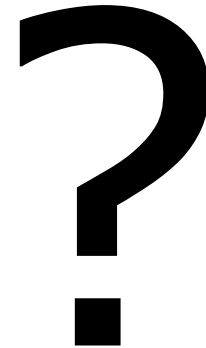
- *imbued with more meaning than their content justifies*
- *viewed as the only statistical outputs needed – value of data for policy making ignored*

Purpose of global cross-nationally comparable indicators

- To provide the global picture
 - ➔ for advocacy
 - ➔ resource mobilisation, engaging donors, demonstrating commitment
 - ➔ accountability of governments – also as *carrots* and *sticks*
- For purposes of comparison
 - ➔ learning from one another - to show what can be achieved
 - ➔ benchmarking
 - ➔ act as a catalyst for debate

The measurability of goals

- Realistic
- Measurable
- Relevant
- Universal
- Coherent



SMART specific, measurable,
achievable, relevant and timed

It is unsmart

- to require that next year's performance is better than current targets and current performance.
- to cascade targets by imposing the same target—for example, that at least 75% of pupils achieve a literacy standard—on a class of 30 as nationally.
- to set an extreme value target, such as 'no patient shall wait in accident and emergency for more than 4 hours' because typically, avoiding extremes consumes disproportionate resources.
- to specify a national target in terms of statistical significance rather than in operational terms.

Goals and even indicators often selected without involvement of statistical expertise

- Inadequate methodological and conceptual development
- Can require significant amounts of data- but additional indicators might be obtained with no requirement for new data
- Presented without metadata (*MDG example*)
- Treated as if error free and context-independent
- Embrace lowest common denominator
- Ambiguity concerning the unit of analysis leading to inappropriate decisions relating to the design of indicator systems
- Often inadequate proxies and over-simplistic

Indicators of development

- over-emphasise the national – distort where we place resources
- pay too little attention to within-country inequalities
- do not reflect variations in priorities/ issues across the world (often turn into normative frameworks)

Indicators to measure change over time

- Assume baseline data
- Introduce rigidity
- Frequency of measurement
- Enough change over time to warrant the monitoring
- Is the change an artefact of noise – essential to know about variation and to understand regression to the mean
- How can we account for changes in expectations or in quality?

Do we measure what we treasure?

*Economists have come to feel
What can't be measured, isn't real.
The truth is always an amount;
Count numbers, only numbers count*

Do we measure what we treasure?

- Focus on financial measures
- Things we can count
- Quality is difficult to measure
- Distortion caused by focussing on one thing
- Focus on comparison not to learn but for the sake of publicity

Problems with league tables

- Lack of appropriate contextualisation – are we comparing like with like?
- Conceptual problems in devising measures of 'value added'
- Input and output measures in different units
- Fail to take account of uncertainties in the data/ uncertainty under-estimated
- Over-interpretation and excessive use of 'stick'
- Inadequate understanding of the limitations
- Often use composite indicators with an inadequate specification of the statistical models

Led to political embarrassment, skepticism

Perverse incentives

- Almost inevitable that some will “play the system” – the greater the incentive, the higher the risk of distortion
- Gaming: generate preoccupation with particular indicators so as to turn attention away from others
- Corner cutting to get the ‘right’ results
- Changing the indicators/ fudging the data/ changing the behaviour
- Hitting the target but missing the point

Unintended consequences

- Effect on staff morale
- Publicly labelling failure
- Creating a blame culture – is naming and shaming the aim?
- Setting yourself up for failure - unrealistic targets

Our responsibilities

- Recognise that monitoring is not a costless activity
- Costs must be proportionate to the benefits
- Be transparent about the limitations of the indicators, and show uncertainty
- ...and about the value judgements which led to them
- Put the indicators in context
- Involve range of stakeholders in selection of goals
- Expose misuse of indicators/ defend integrity of data/ shield from undue political interference
- Understand incentives and manage risks
- Develop protocols?

challenges

- Presentation of indicators in ways which are useful to people who want to make a difference – simplicity v. clarity
- Explore new ways of presenting metadata and especially information on uncertainty, sensitivity analysis
- More research needed on the creation of composite indicators

Aggregating several indicators that purport to measure similar aspects of governance is likely to improve accuracy, on the principle that more information is usually preferable to less.

Learn from good examples

- Some countries have created elaborate and sophisticated program evaluation systems, such as Colombia's National Evaluation System of Public Sector Performance.
- Still, it is difficult to find developing countries that effectively measure the performance of their public administrations in meeting key policy and program objectives, and then use that information continuously to improve policy and program design and performance.
- Even efforts as sophisticated and well-resourced as Colombia's, while producing some impressive performance measures, have failed to impact significantly the quality of policy and program design and implementation.

Learn from good examples

- Chile began a slow but deliberate process of introducing performance indicators into its budget process in 1994, and by 1998 had managed to get both parliamentarians and executive branch leaders more clearly focused on results.
- Importantly, the Chilean effort did not establish performance measures unilaterally by central authorities, nor did it take a "big bang" approach of introducing performance measurement in all agencies and for all policies within a single, short time frame.
- Rather, it created strong budget-process-based incentives for line ministries to develop performance indicators, provided technical assistance to help them do so, and refused to set arbitrary, unrealistic deadlines for the introduction of performance measurement.