

"What I'm determined to do with the cabinet colleagues is to say well, here are some performance benchmarks in terms of how this is implemented over the next three years. .. we do that and everyone's clear about what's expected of them."

Kevin Rudd, interviewed on the ABC's 7-30 Report, 27 November, 2007.

Time only flows one way, but many listeners to Kevin Rudd's first post-election interview with the ABC's Kerry O'Brien must have felt a sense of "déjà vu all over again". Were they in Australia or in early Blairite Britain?

In effect, imposing performance targets on Ministers and their departments was the battle cry of the first Blair government. Elected in May 1997, by 1998 the Blair government had imposed over 300 "headline" performance targets on departments and agencies. Each of those targets – which were linked to budgetary allocations agreed with Treasury – was accompanied by a large number of more detailed performance indicators, with these cascading into yet further targets at lower levels of service delivery. A single "headline" target could translate into 30 or more "measurable outcomes", yielding many thousands of indicators in total.

But while Blair took performance measurement in the public sector to new extremes, he was by no means the first to make large scale use of that approach. Rather, that distinction lies with Robert McNamara's reliance on performance indicators as a way of managing the Vietnam war.

McNamara, who had brought the analytical tools he had developed at the Ford Motor Corporation to the Pentagon, saw no reason why an emphasis on achieving measurable objectives could not improve the performance of his generals, just as it had improved that of Ford's key managers. Out of that belief was born the emphasis on the "body count" as the touchstone of US success or failure in Vietnam.

The consequences that emphasis had highlight the problems with the approach.

The difficulty is that "body counts" are a highly ambiguous, if not very poor, measure of performance in unconventional war. This is because the "body count" tends to increase when the adversary is well-resourced, places a low opportunity cost on lives lost, and hence goes and stays on the offensive. Reflecting this, the "body count" in Vietnam often rose when the adversary was unusually strong, rather than especially weak. An emphasis on the indicator, as against the underlying process generating it, therefore impeded realistic assessment of the situation.

Additionally, and even more seriously, a "body count" target created incentives for resources to be diverted to its achievement. In a war zone, it is not difficult to kill people or count dead bodies. But efforts spent so doing are diverted from tasks that, though less readily quantified, are no less important. The result is to distort the allocation of resources away from what really matters to what is being measured. In Vietnam, this took the form of persistent underinvestment in holding areas that had been cleared, and in making those areas genuinely safe.

Finally, these distortions are increased by the practice known as "storming": that is, achieving the target just before a reporting period ends by throwing resources at the task. Those resources have relatively low marginal productivity in the "big bang" operations to which they have been diverted but create big shortfalls in performance elsewhere. The result is to hit the target but miss the goal.

All of this was clear from the systems analysis assessments made in the course of the Vietnam war and after its end. But, as Karl Marx famously said, history repeats itself, the first time as tragedy, the second as farce. In this case, the Vietnam tragedy was followed by the farce of Blairite managerialism.

Three features dominated that farce.

First, the headline targets came to determine the allocation of effort and resources, regardless of how poorly those targets were related to the objectives actually being pursued. As a result, things were done which would otherwise have been regarded as completely unacceptable. For example, so as to ensure patients presenting at emergency would be seen within the four hour target, some hospitals required patients to wait in ambulances outside, until the queues in emergency had reduced to the point where they could be admitted without putting the target (rather than themselves) at risk.

Second, the focus on the targets induced evasion, including through manipulation of data. Just as Soviet managers routinely lied about whether they had achieved production quotas, so "creative accounting"

became widespread, including by altering measurement rules to exclude cases that would have led to the targets not being met.

Third, as embarrassments mounted, the system was corrupted to make it less politically troublesome. Targets were made softer; the number of indicators was multiplied to the point where it was impossible to understand what was being measured or achieved; and little effort was put into auditing glowing reports from the field.

Reviewing these outcomes, the Royal Statistical Society concluded that while performance monitoring "is broadly productive" when done well, "done badly, it can be very costly and not merely ineffective but harmful and indeed destructive".

But the Blairite experience holds a deeper lesson as well. No matter how much thought goes into performance measurement, central planning is central planning, with all of its pathologies. True, firms in the private sector can and do invest in designing better systems for setting goals and tracking performance; but what really drives their performance is the thirst for profits and the threat of displacement, as the loss of custom to competitors provides the clearest signal of the need to lift one's game.

In short, an ounce of market incentives is worth many pounds of KPIs. If the Rudd government wants better performance from public providers of health, education and community services, it would do well to pursue its predecessor's emphasis on exposing them to competition and choice. That is not because markets are perfect – far from it. But the performance indicator that can do better remains to be invented.