



How Frequent Reporting of Quantitative Accountability Measures Can Undermine Bureaucratic Performance

Dan Honig, Georgetown University

There is a long history of measurement misleading policymakers. Current National Security Adviser H.R. McMaster's account of the Vietnam War, *Dereliction of Duty*, suggests that an over-reliance on quantitative measures hurts military performance, shifting control from field personnel to analysts based in headquarters. Some postmortems of Hillary Clinton's 2016 campaign find fault in its privileging of data-driven planning from headquarters over the judgments of workers in the field. Obviously, data can be helpful, but when and where do top-down, data-driven techniques lead to better performance and when do they hinder it?

This question applies to attempts often made by legislators to improve the functioning of bureaucracy by setting performance targets and measuring progress against those targets. In the United States, for example, the Government Performance and Results Act has required agencies to file quarterly reports since 1993 assessing their performance "in an objective, quantifiable, and measurable form." My research finds that the push to meet targets can sometimes undermine rather than improve performance. I also explore the conditions allowing for more or less success.

When Reporting Undermines Performance – Lessons from Foreign Aid

Aid agencies such as the World Bank, U.S. Agency for International Development, and the United Kingdom's Department for International Development provide a useful setting for understanding the effects of management on performance across tasks and environments. The range of foreign aid efforts is vast; aid agencies are involved in everything from education, health, and infrastructure policy to public financial management, anti-corruption efforts, and judicial reform. Indeed, it is difficult to come up with a public task or sector from which aid agencies are entirely absent. By comparing the functioning of the same agency across countries and tasks as well as the functioning of different agencies doing similar tasks, we can learn much about the circumstances in which top-down accountability and performance measurement are more or less beneficial to agency performance.

What kind of measurement can be employed when the purpose of an activity is not to deliver concrete, measurable objects but, for example, skills and expertise? The standard measure used by the U.S. Agency for International Development for this kind of activity involves counting the number of individuals trained for some task, with the total reported back to U.S. Congress as evidence of success in the activities. But of course, gaining expertise is not as simple as attending a training session, so the numbers of people trained are not direct evidence of accumulation of necessary skills. What is more, this approach to measurement gives those delivering and supervising the training every incentive to focus on maximizing the number of people trained regardless of whether they are the appropriate individuals or are actually going to use the skills. In this type of situation typical in development efforts, the push for immediate data to monitor performance actually hinders

rather than helps the accomplishment of laudable objectives.

Drawing on case studies plus a database of development project outcomes I have assembled – with over 14,000 projects over 40 years – my research finds that aid agencies achieve better outcomes when given scope for independent action rather than being tightly supervised using narrow measures. More agency autonomy translates into more personnel empowered to act effectively in their assigned countries. The less frequently individuals on the ground are required to meet quantitative targets, the more creative and less hamstrung they will be – taking smart risks, rather than acting so cautiously as to make sure they never err.

The value of agencies able to act independently rises when things get messy. In more fragile states and in projects with hard-to-measure outcomes, flexible agencies are more adept. This finding is consistent with other research on organizational effectiveness. From factory workers to loan officers, overly close supervision can discourage workers from making hard to justify judgments that would lead to better organizational performance. The lesson is clear: intuition and on-the-ground expertise can make a real difference.

What Determines the Effect of Top-Down Performance Measurement?

Performance measurements and top-down control in general are sometimes very useful. Two key factors determine whether they are likely to work well, according to my research:

- To what degree are unexpected developments – “unknown unknowns” – likely to influence outcomes? Tight control may be more useful in Kansas than it is in Kabul.
- How amenable are tasks to monitoring and measurement? When delivering medication or constructing a road, tighter control is associated with better outcomes. But this is not as true for tasks where success depends on changing circumstances or hard to quantify “judgment calls”.

What Can Policymakers Do?

Top-down controls, including monitoring and measurement, are tools – not “good” or “bad” any more than a screwdriver is “good” or “bad.” Where these tools are not inappropriate, there are a number of accountability strategies managers and political authorizers can employ instead. Qualitative performance reviews can help those in charge understand which individuals and units are most effective. Teams can be placed within an agency to understand what may be holding back performance. In some cases, reporting can be more helpful if the time period is customized according to the task, rather than, for example, requiring reports quarterly in all agencies.

When trying to improve bureaucratic performance, policymakers should think carefully about the complex effects of measurement and control. Slapping targets and reporting requirements onto an under-performing agency may lead to more paperwork without better performance. Managing by numbers works well for well-defined tasks – such as vaccine delivery – and should be applied more often for such endeavors. But for some tasks policymakers should keep in mind that rigid, numerical reporting routines can do more harm than good. This is especially likely to be true when organizational agents need experience and practice at making complex judgements and on-the-spot-assessments. For such tasks, more holistic and flexible forms of assessment are needed.

