

# Project Syndicate

## Long Reads

# The Metric God That Failed

Mar 16, 2018 | **JERRY Z. MULLER**

SILVER SPRING, MARYLAND – In 1986, the American management guru Tom Peters popularized the organizational theorist Mason Haire’s dictum that, “What gets measured gets done,” and with it a credo of measured performance that I call “metric fixation.” In time, the devotees of measured performance would arrive at a naive article of faith that is nonetheless appealing for its mix of optimism and scientism: “Anything that can be measured can be improved.”

In the intervening decades, this faith-based conceit has developed into a dogma about the relationship between measurement and performance. Evangelists of “disruption” and “best practices” have carried the new gospel to ever more distant shores. If you work in health care, education, policing, or the civil service, you have probably been subjected to the policies and practices wrought by metric-centrism.

There are three tenets to the metrical canon. The first holds that it is both possible and desirable to replace judgment – acquired through personal experience and talent – with numerical indicators of comparative performance based on standardized data. Second, making such metrics public and transparent ensures that institutions are held accountable. And, third, the best way to motivate people within organizations is to attach monetary or reputational rewards and penalties to their measured performance.

This credo is attractive to politicians, policymakers, corporate directors, and non-profit trustees, because it marries good intentions to managerial technique. With “hard” numbers, you can justify the use of “hard” incentives based on cash or reputational accolades. Accordingly, the trinitarian formula “measure, monitor, and remunerate” has come to be regarded as a panacea for even the most difficult social problems.

Yet the outcomes of metric-driven approaches in health care, education, and public safety

have been disappointing – even counterproductive – and for similar reasons across sectors. We tend to attribute the success or failure of individual hospitals, schools, police forces, and other formal institutions to the institutions themselves. But what matters more is the broader social and demographic environment in which they operate.

## **Metrics, Heal Thyself**

Consider the US medical system, which comprises clinics, hospitals, government agencies, private insurers, and millions of physicians and other health professionals. According to the World Health Organization's oft-cited *World Health Report 2000 - Health Systems: Improving Performance*, "It is hard to ignore that ... the United States was number 1 in terms of healthcare spending per capita but ranked 39th for infant mortality, 43rd for adult female mortality, 42nd for adult male mortality, and 36th for life expectancy."

In fact, by international standards, the US medical system is highly successful in diagnosing and treating disease. But mortality and life expectancy are largely influenced by factors *outside* the medical system, not least those relating to culture and lifestyles. On average, Americans are more obese than other nationalities (though some are catching up quickly), and it is well known that obesity is associated with chronic and debilitating illnesses such as Type-2 diabetes and heart disease.

Similarly, by international standards, Americans were heavy smokers up through the 1980s, and smoking contributes powerfully to heart disease, cancer, and other ailments, possibly even decades after a person quits. The US also has disproportionately high rates of death from gunshot wounds – a tragic fact that has almost nothing to do with the medical system.

In short, America owes many of its leading health problems not to medical-system failures, but to independent social and cultural factors. National metrics of mortality and longevity depend heavily on whether people get enough exercise, eat right, handle firearms responsibly, and refrain from unsafe sex, smoking, and alcohol and drug abuse. Though physicians and public health officials can and do try to influence these habits, they ultimately have little to no control over individual lifestyle choices.

## **Eat Your Carrots (and Sticks)**

The metric-driven approach has also fallen short at the level of medical institutions. According to the canon's second tenet, hospitals' performance can be improved through "transparency" and "accountability," by making all hospital mortality rates public. Yet a recent study in the *Annals of Internal Medicine* found that "public reporting of mortality rates has had no impact on patient outcomes."

In 2009, Medicare began demanding that all acute-care hospitals report their readmission rates for patients treated for major medical conditions, and it published these data on its "Hospital Compare" website. Then, in 2012, it went from public reporting to paying for performance, by imposing financial penalties on hospitals with above-average readmission rates. The idea was to incentivize hospitals to reduce readmissions, and thereby costs. Yet, as

of 2015, three-quarters of all reporting hospitals were still subject to Medicare penalties.

Worse still, the hospitals most likely to be penalized were those serving the most difficult patient populations. For example, major teaching hospitals were disproportionately affected, because they tend to see patients with more problematic conditions. So, too, were hospitals in poverty-stricken areas, where patients often lack access to care, or are unable to care for themselves, after being discharged from the hospital.

Again, we find that the determinants of relative success and failure lie beyond the bounds of the institutions being measured, rewarded, and penalized. And though pay-for-performance schemes try to compensate for this perverse outcome through a process known as “risk adjustment,” calculations of risk, like any other metric, are prone to misspecification and manipulation.

Moreover, the failure of pay-for-performance schemes in medicine is not confined to the US. According to a sophisticated, comparative 2016 study published in *The Lancet*, the British National Health Service’s “Quality and Outcomes Framework,” the largest pay-for-performance primary-care program in the world, has had no significant effect on mortality rates in the United Kingdom.

### **Learning by the Numbers**

The metric fixation has also been the bane of primary and secondary schooling in recent decades. Under a succession of US presidents from both major parties, the US Department of Education and state governments have poured resources into standardized-testing regimes that reward or punish teachers and schools on the basis of published results.

Yet, as Daniel Koretz of the Harvard School of Education shows in his recent book *The Testing Charade: Pretending to Make Schools Better*, this metric-based approach has had a minimal net effect on student learning. Meanwhile, it has had a demoralizing impact on teachers, who are vilified as lazy or incompetent when the measured performance of their students is found wanting.

In reality, schools have only a limited effect on relative levels of educational attainment – and that should not come as a surprise. As we have known at least since the 1966 “Equal Educational Opportunity Study,” which was commissioned by US President Lyndon B. Johnson’s administration, student performance is closely correlated with the social, economic, and educational attainment of the parents.

Despite the demonstrable failure of pay-for-performance schemes in education, the federal government has continued to dump ever more resources into such efforts. And here, too, the US is not alone. The UK, Portugal, Australia, Chile, Mexico, Israel, and India have also taken steps to link teacher pay, tenure, and promotion to students’ measured performance.

### **Metric-ulation Nation**

The metric evangelists have also set their sights on higher education. Organizations such as the Bill & Melinda Gates Foundation and the Lumina Foundation have encouraged more people to attend college and urged state governments to reward and punish state colleges on the basis of graduation rates. And, over time, more and more Americans have entered college after graduating from high school.

There is no indication, however, that college preparedness in the US has increased alongside attendance rates. One measure of preparedness is a student's performance on aptitude tests such as the SAT and the ACT. The company that markets the ACT has developed benchmarks to indicate whether a test-taker has a "strong readiness for college course work." Of those who took the ACT most recently, one-third did not meet that benchmark in *any* of the four areas tested (English, math, reading, and science), and only 38% did so in at least three.

These results suggest that many of those who aspire to go to college do not have the demonstrated ability to succeed once they are there. In fact, because many of the students entering two- or four-year colleges are unprepared, they often must enroll in what are euphemistically known as "developmental" courses covering high-school-level material. As of 2010, one-third of students entering two-year community colleges were placed in developmental reading courses, and 59% were referred to developmental mathematics courses.

One predictable consequence of colleges' influx of unprepared students is a decline in graduation rates – a widespread phenomenon that imposes substantial costs on students in the form of tuition, living expenses, and foregone earnings. And this comes at the same time that state governments have been increasing the pressure on public colleges to boost their graduation rates.

But, again, college completion depends less on teaching performance than on the aptitude of incoming students. So, in response to the pressure from legislators, some public colleges have been quietly lowering the bar for graduation. In practice, this means offering more courses with easily fulfilled requirements, and compelling professors – sometimes overtly, sometimes tacitly – to inflate grades. This process is facilitated by the increasing use of adjunct instructors, whose contracts need not be renewed if they fail too many students.

Such metric-oriented practices allow college administrators to point to "transparent" graduation-rate data to demonstrate a high level of performance. What is not transparent is the manner by which standards have been lowered. As in health care, the formal institutions of higher education are blamed for factors outside of their control, and efforts to game the system are rewarded.

### **Arrested Development**

A final area in the grip of our era's mania for measurement is policing. When voters and politicians think of public safety, they think of the police, whom they hold responsible for

crime rates. Yet, like health and its relationship to the medical system, or education and its relationship to the school system, public safety is only partly dependent on effective policing.

After all, public safety also depends on other elements of the justice system, such as prosecutors, the judiciary, and the penal and parole systems. It depends on the local population's propensity for criminal activity, which in turn depends on broader socioeconomic and cultural factors. And it depends on whether there are opportunities to commit crimes. Indeed, part of the dramatic decline in US crime rates in recent decades resulted from steps taken by private property owners, not the police. The proliferation and increased sophistication of car- and home-alarm systems has made theft and burglary far more difficult. And some one million Americans are now employed by private security firms.

As with other formal institutions that have been subjected to metric fixation, police forces have learned – indeed, been encouraged – to game the system. To “reduce” crime rates, they can book felonies as misdemeanors, or not book arrested suspects at all.

Politicians will always be tempted to impose unrealistically high expectations on institutions to show that they are looking out for the public good. It is nice to think that with the right incentives in place, everyone can be healthy, all students can succeed, and every city can be crime-free. And it is inevitable that a cottage industry of organizational gurus and international consultants would capitalize on these hopes by peddling “best practices” and metric cures to deep-seated social problems. But, as we have seen, the outcomes are often disappointing.

This is not to say that data-driven measurement of performance is useless. Performance metrics can be beneficial, especially when used not for reward and punishment, but for diagnostic purposes by the practitioners themselves, and are shaped by their professional values, experiences, and expectations. But, all too often, metric fixation creeps in. The metrical canon promises a solution to any problem, no matter how intractable. And if a solution does not materialize, it can always produce a sacrificial lamb to bear the blame and exorcise uncertainty.



## JERRY Z. MULLER

Writing for PS since 2003

3 Commentaries

Jerry Z. Muller, a professor of history at the Catholic University of America, is the author of *The Tyranny of Metrics*.

<http://prosyn.org/he0JVKW>;

---

© Project Syndicate - 2018