

The Tyranny of Metrics

Jerry Z. Muller. 2018. Princeton: Princeton University Press. ISBN 978-0691174952. \$24.95 (hardcover).

Reviewed by Jamin Hübner, The University of the People

It is difficult to overstate the influential role of metrics in today's world. Whether in sliding our finger across a screen to buy coffee downtown, or scheduling a doctor's appointment for a checkup, our experience is largely determined by analytical goals working in the background. In theory, all of this data collection and goal-setting is meant to increase performance—and we tend to assume that it ultimately will.

Historian Jerry Muller thinks otherwise. He contends in *The Tyranny of Metrics* that not only is the contemporary reliance and enforcement of metrics based on flawed premises, but results in adverse unintended consequences. His thesis is relatively simple, and it highlights a problem experienced by anyone who has wrestled with the value of data collection and performance goals: “There are things that can be measured. There are things that are worth measuring. But what can be measured is not always what is worth measuring; what gets measured may have no relationship to what we really want to know. The cost of measuring may be greater than the benefits” (p. 3).

Modernism's philosophical hegemony and its obsession with the quantifiable, propositional, and “objective” has more or less given rise to what Muller calls “metric fixation.” “The most characteristic feature of metric fixation,” he explains, “is the aspiration to replace judgment based on experience with standardized measurement. For judgment is understood as personal, subjective, and self-interested. Metrics, in contrast, are supposed to provide information that is hard and objective” (p. 7). The trouble, however, is that this framework is hopelessly naïve. Judgment cannot always be substituted, and numbers can be misleading, manufactured, misinterpreted, and/or improperly implemented in a variety of ways. As the rest of the book argues, *such numbers frequently are*.

Chapters seven through thirteen look at case studies of how the use of metrics leads to unintended consequences. In colleges and universities, accreditors' demand for “measurable outcomes” privileges certain academic programs over others (p. 87) and distracts faculty from educating, forcing them to instead climb over often irrelevant bars. The classic benchmark of high graduation rates actually undermines

other goals like accessibility, “since less advantaged students tend to be not only financially poorer but also worse prepared” (p. 83).

In primary and secondary schools, the consequences are more discouraging. After surveying this notorious subject in American history, Muller concludes that “General improvements in schooling do not therefore lead to greater equality of outcomes” (p. 98). More specifically, “Improvements in the quality of schools may elevate overall educational outcomes, but they tend to increase, rather than diminish, the gap in achievement between children from families with different levels of human capital” (p. 98). Efforts like the No Child Left Behind Act have only amplified negative consequences and placed teachers in an unwinnable game of “demoralization” (p. 100).

The medical field also presents serious concern. Surgery success rates encourage physicians to avoid high-risk cases that might damage their reputation (and their institution’s reputation), leaving many patients without proper care and (sometimes) left to die. Hospitals, likewise, game the system by manipulating the status of readmitted patients (p. 121).

Similar phenomena are witnessed in the notorious cases of “crime rates” (police), body counts and SIGACTS (violent activity in the military), the endless abuses of metrics in the business world, and distortions created in nonprofits by seeking certain metrics and criteria by which foreign aid is issued. The problems in all of these cases are the same—gaming the system, cheating, and manipulation.

All of these case studies also reveal the subtle and sobering reality: quantitative benchmarks can be used to *mask poor* performance as much as they can be used to *achieve good* performance. Creating targets simultaneously create new “hazards” because of the way they “influence behavior” (p. 87).

In my reading of the book, this is really the big philosophical conundrum lurking behind the issue: metrics are meant to put under control something that was not under control. But creating a *tool for control* (e.g., metrics) is not inherently positive, and at times can result in new negative consequences that were not a problem to begin with. Metrics are like knives. A few necessary and properly-used tools can completely change the experience and functionality of food-making in a kitchen. But it would be foolish to thereby buy as many knives as possible and stack them all around the house in hopes of achieving greater

“performance.”¹ (Hence the book’s concluding chapter: “When and How to Use Metrics: A Checklist”).

To put the problem in economic language, the law of diminishing returns applies to the production and use of metrics just like any other experience in the world.² A few key metrics can massively improve the performance of a business or one’s knowledge about entire economies. Muller does not (and could not) deny this. But, as another law of economics states, “there is no free lunch.” The utilization of metrics comes at a cost. Many managers, leaders, and entrepreneurs behave as if this was not the case—or act as if the only cost is a new analyst.

This general argument of the book is supported by the social sciences. “Campbell’s Law,” for instance, states that “[t]he 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” (cited on p. 19). Similarly, Goodhart’s Law states that “Any measure used for control is unreliable” (p. 20). Again, quantifying phenomena is not positive in and of itself, and only creates another layer of organization capable of corruption and misuse.

Standardization also degrades information. In Muller’s words, “... making things comparable often means that they are stripped of their context, history, and meaning” (p. 24). This creates a false sense of truth and certainty; “information appears more certain and authoritative than is actually the case: the caveats, the ambiguities, and uncertainties are peeled away, and nothing does more to create the appearance of certain knowledge than expressing it in numerical form” (p. 24). In a cultural context that privileges the “scientific” and “measurable” by default, this potentially deceptive aspect of metrics is also amplified.

One can see, then, that Muller’s argument is more sophisticated than it initially appears. I also found it interesting how he connects America’s litigious ethos with quantifying human experience; “the *perception* of the United States as a litigious society creates an anxiety about the possibility of being sued, leading to defensiveness and risk-aversion” (p. 42). Managers and leaders complain about too many lawyers while creating the very kind of culture that needs them.

I came to *The Tyranny of Metrics* with enthusiasm and was not disappointed. This might be due to the fact that I also came to the book with an unusual bias. I remember one of my favorite college professors

getting laid off—the number one organist in the world at the time—because the institution needed to achieve the “golden student-to faculty-ratio” in order to maintain its place in the top ten colleges of *US World and Report*.³ For the past four years, I served as the chief accreditation officer to bring a startup college from applicant to accredited status and struggled regularly with creating assessment reports and ensuring that the college was functioning according to “measurable outcomes.” My wife Jessica has watched the world of metrics destroy the field of mental health as the medical model of psychiatry—and its marriage to the insurance industry—force counselors to mechanically diagnose patients and act according to an “evidence-based model” of therapy. Her father, at one point a top-producing foot-surgeon in the U.S., has told endless stories about how metrics and regulations in the medical industry (and VA system) have unintended consequences. (For many years, I’ve therefore considered writing a book entitled *Evidence-Based Nonsense*. Fortunately, it looks like this kind of book is already written!)

Muller’s *Tyranny of Metrics* was not flawless, of course. For many readers, the whole problem doesn’t need identifying or arguing. “Metric fixation” could be viewed as just an extension of bad econometric and statistical methodology. Some chapters were also less than five pages long. And some chapters could have been added—such as the aforementioned problem in mental health, where “metric fixation” and the insatiable lust to quantify compromises the entire field.⁴

Muller also expressed some somewhat disconnected and dubious opinions. In one section in the book, he ventures out on a trail about the problems of public transparency, railing against Wikileaks and Edward Snowden. Whatever relevance this may have, it’s not clear that the arguments are sound. We read that “Snowden was hailed as a hero by portions of the public in the United States and Europe. At the heart of Snowden debacle lies the belief that transparency is always desirable” (p. 163). But many interpret the situation differently; the “heart” of the “debacle” was that one of the most powerful governments in the world was (at least on some issues) functionally accountable to no one, and sometimes there are no good options when major ethical lines have been crossed.⁵

Fortunately, these are minor issues. Overall, *The Tyranny of Metrics* is timely and couldn’t be more relevant to so many of today’s spheres of society.⁶ It is very easy to read and would make for good supplementary

reading in a statistical, econometric modeling, or sociology class. “Ultimately, the issue is not one of metrics versus judgment, but metrics as informing judgment, which includes knowing how much weight to give to metrics, recognizing their characteristic distortions, and appreciating what can’t be measured. In recent decades, too many politicians, business leaders, policymakers, and academic officials have lost sight of that” (p. 183).

Endnotes

1. This is my analogy, not Muller’s.
2. Although Muller does not explicitly put it this way, he points in this direction (see pp. 3 and 178).
3. At the time, this Golden Ratio for liberal arts institutions was 15:1. I am unaware if it has changed for current trends.
4. The work of Michael White and David Epston have pointed out this unfortunate trend for many decades. Cf. Robert Whitaker and Lisa Cosgrove, *Psychiatry Under the Influence: Institutional Corruption, Social Injury, and Prescriptions for Reform* (New York: Palgrave MacMillan, 2015).
5. Similarly, Muller says “‘Sunlight’ is the best disinfectant’ has become the credo of the new faith of Wikileaks: the belief that making public the internal deliberations of all organizations and governments will make the world a better place” (p. 164). This is, in my opinion, to confuse the purpose of Wikileaks with some of its exceptional abuses. Again, some governments and states only give citizens bad and worse options of response.
6. Princeton University Press deserves special commendation for their consistent signature binding (a rare but pleasant feature in a world of monograph publishing that seems fixated on perfect bound), especially for a book of short size.

Bibliography

Whitaker, R., & Cosgrove, L. (2015). *Psychiatry Under the Influence: Institutional Corruption, Social Injury, and Prescriptions for Reform*. New York: Palgrave MacMillan. ■