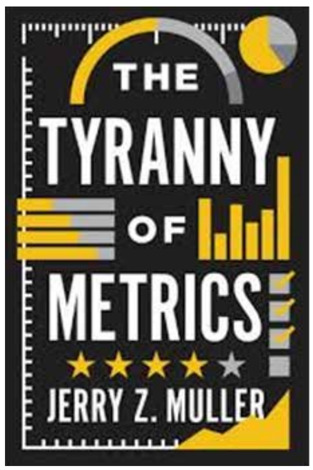
The Tyranny of Metrics

In his book, ***The Tyranny of Metrics***, Jerry Z. Muller takes aim at the widespread practice of employing standardized metrics to track and manage the performance of organizations and networks working on complex issues.



Muller is not blindly anti-metric. He points out that quantitative data can help decision-makers get a better understanding of the challenges we are trying to address (e.g., big data to show where crime is likely to happen) and be help track patterns and trend over time (e.g., How well are kids reading in grade 3 over the last 5 years?). His concern is that pure numerical evaluations of progress often do not lead to better outcomes, and that western management requires a radical rethink about how to approach assessment.

The core of Muller's argument is that our collective fixation with metrics is based three shaky beliefs:

- That outcomes and performance can be reduced to standardized measurements even in complex challenges, to the neglect of qualitative data and the interpretations and judgments of people with experience and expertise in these matters.
- That making metrics public or transparent increases the chances that institutions will be accountable for their actions and results.
- The best way to motivate people is by attaching financial and reputational reward and penalties to their ability to meet quantifiable performance targets.

Muller puts each of these beliefs to the test by drawing on an impressive mix of studies, anecdotes, and expert opinion in diverse areas of education, medicine, policing, foreign aid, the military, business and finance, and philanthropy.

His conclusion is simple. Not only does a reliance on standardized metrics distort our view of reality, it leads people to adopt perverse behaviours. Some of the most common ones include (a) relying on metrics that are easy to collect in the short term, even if they reveal little about the real outcomes people are pursuing; (b) focusing on those metrics that are most closely monitored at the expense of other parts of the work that may turn out to be equally important; and (c) producing inaccurate data to demonstrate compliance and/or progress towards an objective.

So predictable is the gaming of performance metrics, that the evaluation pioneer Donald Campbell crafted a law to capture the phenomenon: "The 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."

Despite these profound and now obvious challenges, decision-makers seem unable to shake their Obsessive Measurement Disorder. Staff frustrated with the limitations of their first batch of performance metrics, dutifully produce a second batch, in the vain hope that the next one will yield better insights. Even organizations that have given up hope of finding useful metrics continue to gather and report them as a way of "virtue signaling," i.e., demonstrating their commitment to "measure something," even if the data they do collect largely gets ignored. The costs of these well-meaning, yet unproductive behaviors are significant: lost time and resource, falling morale, bad decisions, and weaker performance.

While Muller is fierce in his criticism of the mainstream approach to employing metrics, he is careful to point again and again that it's not the act of measuring that creates problems, it's the inappropriate, excessive, and simplistic application of standardized metrics to decision-making. The solution, he argues, is to treat metrics as one part of a larger evaluation process that draws on multiple types of data and invests appropriate time and expertise to interpret that data.

To demonstrate this, he points to a controversial, but highly instructive case study: the efforts of NATO partners to get a handle on the extent to which their efforts are leading to peace and security in Afghanistan. Here central military administrators and politicians agree that the two indicators of counter-insurgency campaigns preferred by high-level commanders, the media, and public officials, while easy to measure and aggregate, are inadequate:

- The number of "enemy fatalities," a metric that does not offer insight into enemy's will to fight, the central factor in such conflicts.
- Decreases in violent incidents in a region, which does not reveal if true peace and stability has been achieved, or if insurgents are quietly in control of the area.

To get a more meaningful understanding of progress, front-line commanders have come to rely on more nuanced, locally-specific indicators of stability. This includes the price of bananas in isolated regional markets. He quotes

> Afghanistan is an agricultural economy, and crop diversity varies markedly across the country. Given the free-market economics of agricultural production in Afghanistan, risk and cost factors – the opportunity costs of growing a crop, the risk of growing it at market and of transporting money home again – tend to be automatically priced in to the cost of fruits and vegetables. Thus, fluctuations in overall market prices may be a surrogate metric for general popular confidence and perceived security. In particular, exotic vegetables – those grown outside a particular district that have to be transported further at greater risk in order to be sold in that district – can be a useful telltale marker. (Muller, p. 135).

Ben Connable, an expert in counter-insurgency for the Rand Corporation, has researched and built on these spontaneous evaluation practices to propose a model of "contextualist assessment." The approach has several key features:

- Allow front-line units to employ their own, context specific, metrics, in addition those required for centralized reporting, so that they can have meaningful feedback on their local area.
- Encourage local commanders to complement standardized metrics for central use with select local indicators of stability, along with their own interpretation of the state of affairs in their zone of operation, in order to give more context and meaning for whatever hard data they share.
- Employ strong processes for making sense of data at the centralized level, involving people with experience and judgement, in order to understand what diverse sets of data are saying about current and future stability.

This concept of contextual assessment is instructive to any effort to tackle a complex issue. Take, for example, the widespread use of the annual "homelessness count" in many Canadian cities. Social service agency staff and volunteers spread out over the community spot, interview, and record the number of people living on the streets. While most people involved in homelessness work agree that the numbers produced through these studies shed ***some*** light on the state of the problem, many point out critical flaws in its methodology, and argue that it's too simplistic to be a barometer of success. In order to get a more accurate picture of what is going on, they argue, decision-makers have to draw on a broader range of data and analysis: for example, the number of agency clients without a home address; the unique ways in which homelessness manifests in different parts of the city (e.g., couch-surfing or sleeping rough by the river); and front-line workers' assessments of the complexity of the challenges faced by the people being counted.

Here again the lesson is repeated. The problem is not the homelessness count. The problem is relying on one standardized and centralized metric to provide a simple picture of a complex situation.

The Tyranny of Metrics is an imperfect book. Because it covers so much ground in 180 pages, it not always able to explore important ideas in depth. It would have been useful, for example, to introduce the reader to the work of James Q. Wilson (1989). Wilson describes four types of public service, each of which requires a different approach to measurement and accountability: production services (e.g., passport provision); procedural services (e.g., prison management); craft services (e.g., policing), and coping services (e.g., social work). This simple framework would have alerted readers to differences of contexts and the issue of which metrics (if any) are applicable to each. Muller does not have the time and space to reach this level of analysis.

Muller does not extend a lot of advice about how to avoid the pitfalls he identifies, either. The checklist in the last chapter describes ten things to consider when exploring if and how to employ metrics. These provide an excellent summary of the book's main points and can be used to encourage enthusiastic advocates of measurement to proceed with caution. Yet, readers who want more practical guidance on the nuts and bolts of crafting good metrics will have to find it elsewhere.

These limitations aside, *Tyranny of Metrics* is valuable resource for anyone tasked with producing or using metrics in their work. It's an accessible and compelling account of why our society is so metric-obsessed and the damage that the obsession yields. It also offers some very useful hints about how we correct our approaches so that measurement helps, rather than hinders, our efforts to make a better world.

Learn More

CBC News. November 30, 2016. Edmonton Homeless Count Misleading, Inner City Agency Says. Retrieved from: https://www.cbc.ca/news/canada/edmonton/homeless-count-boyle-street-criticism-1.3874422

Connable, Ben. 2012. Embracing the Fog of War: Assessment and Metrics in Counterinsurgency. National Defense Research Institute. Santa Monica, CA: Rand Corporation.

Muller, Jerry Z. 2018. The Tyranny of Metrics. Princeton, NJ: Princeton University Press.

Wilson, James Q. 1989. What Government Agencies Do and Why They Do It. New York, NY: Basic Books.

Resource Reviews are a series of documents exploring new frameworks, tools and resources for building communities and solving tough challenges. This particular document was developed in cooperation with Tamarack Institute as part of its efforts to build capacity for community change makers.

http://here2there.ca/

http://www.tamarackcommunity.ca/