OBJECTIVITY LAUNDERING:
THE TWO MASKS OF METRICS
C. Thi Nguyen

Metrics hit us on two fronts. First, they discourage reflection and exploration by creating a *facade of objectivity*. Reflective control involves making a choice, but metrics disguise and rewrite the situation so that we don't realize there was a space for choice. Metrics try to speak with the voice of scientific authority; they make it seem like there is no room for dissent. Second, metrics offer us the pleasures of *seductive clarity*. They make things look so simple that we won't think there is anything else to investigate. They present themselves as the bottom line—as the final word, as a finished thought. These are the two masks of metrics, which conceal our freedom to choose: the mask of fake objectivity and the mask of seductive clarity.

Let's start with the mask of fake objectivity.

Theodore Porter, the historian of quantification, wanted to understand why bureaucrats and politicians seem to obsessively and unthinkingly reach for quantitative justification. Administrators often seem to prefer metrics to any other form of justification, even when those metrics are clearly flawed or incomplete. This is because, says Porter, metrics gave the appearance of objectivity—which lets their users avoid taking responsibility. Metrics let administrators avoid visibly inserting themselves into the stream of decision-making so that instead, they can point to the metric and say, "It's not me, it's just the numbers."

But this objectivity is only a facade. Such metrics often contain value judgments hidden at the core. We take a subjective choice and then hide it under tons of precise math. It's like money laundering, where we take dirty money and then pass it through enough clean transactions to mask the dirt.

Let's call this *objectivity laundering*. We take a complex matter, like well-being, education, or success. Somebody—often, a very distant somebody—makes a value-laden decision about what that means, about what counts as well-being or success. Then we process it. What comes out the other end looks objective and free of any taint of human values. It seems to speak with the voice of God—or at least the voice of science.

This is the first way that metrics worm their way into our hearts: by presenting themselves as simple matters of fact. They try to convince us that we never had a choice about the matter in the first place. The more universal and objective that metric looks, the more we are tempted to treat it as a given.

I'm worried that in many of our cases, we're playing the role of the responsibility-avoiding bureaucrat in our own lives. We're hiding from the complexity of our value decisions by following an apparently objective metric. But really, we are deferring to the values hidden within that metric. The existentialist philosophers had a term for this: acting in *bad faith*. What they meant was that we humans have the freedom to choose, but sometimes we try to avoid that freedom. We deceive ourselves into thinking that we have no choice in the matter, that we are acting in the only way possible. We are pretending to be mechanical objects. Jean-Paul Sartre thought we often did this by following the scripts attached to our roles and social identities—letting ourselves think that we had no choice because we are a parent, or a police officer, or an American, and that's simply what we had to do.

Value capture is the new bad faith.

There are different strategies for objectivity laundering. The first is *burying the values*. This happens when we take a value-laden decision and bury it under a mountain of objective mechanical processing, so that we lose sight of the original value-laden choice.

Historian Mary Poovey gives a fantastic example of burying the values, from early in the history of quantification culture. She was interested in understanding the history of something she called "the modern fact" in Western civilization. Before the modern era, she says, we had a very different idea of what made a good fact. We trusted a fact if it came from a trustworthy person—somebody honest, reliable, and competent. Good facts came from good people.

But somewhere on the road to our contemporary world, things changed. We started using a new conception of a fact, a new standard for trustworthiness. A fact was good because it came from *nobody in particular*. It was trustworthy if it came not from a good person, but via a mechanical process. Facts were trustworthy if nothing human had tainted them. They came either from a literal machine or from people who were following clear mechanical rules, which made them into the functional equivalent.

Metrics try to speak with the voice of scientific authority; they make it seem like there is no room for dissent.

This is the idea behind the modern practice of science and data collection. But, says Poovey, the idea of the inhuman modern fact actually shows up a little bit before the birth of the sciences—in the sixteenth-century invention of double-entry bookkeeping.

Double-entry bookkeeping is a methodology, a precise set of clear and mechanical rules for processing the numbers in your accounting books and double-checking your calculations. Since the rules are mechanical, anybody can do the math, and anybody else can check over the accounting process. Accounting, says Poovey, is the first place where we see moral virtue and human competence displaced from the center of our ideal of human knowledge and replaced with trust in a mechanical system. The precision and transparency of the system is what lends credibility to the bookkeeping procedure. Accounting, it turns out, is the birthplace of our modern conception of objectivity—in which human virtue is replaced with mechanical precision. (And Poovey finds direct evidence that the early innovators of science knew of, admired, and actively imitated the methods of double-entry bookkeeping.)

But that mechanical objectivity is incomplete, says Poovey. There is, hidden at the very root, an entirely nonmechanical, non-inspectable procedure: how we set the starting values. Let's say you're a merchant with an inventory. When you start using this system of double-entry bookkeeping, you need to enter a value into your books for each item in your inventory. Once you've entered those values, the wholly mechanical system takes over. But there are no mechanical rules for the initial valuation. That can come from anywhere—intuition, experience, guesswork, bias, fabrication. But despite that raw, uncontrolled core element, we trust the books because the rest of it is so clear. The subjectivity of the initial input is concealed by the mechanical objectivity of the ensuing processing.

This kind of objectivity laundering remains very common. Economic cost-benefit analyses usually depend on some obscured but crucial value decisions inserted quietly into the calculations. Here's an example from Porter: The National Park Service often uses cost-benefit analysis to justify investments in infrastructure. It's relatively easy to calculate, in a fairly objective manner, the projected profits for local businesses from increased tourist traffic. But the National Park Service also has to calculate the *value of visiting the park* for the tourists themselves—to attach an economic value to enjoying nature. So they set one. One 1948 cost-benefit analysis set the "recreational value" of each tourist visit at 12.5 cents per visitor day. Recreation is a key value of the national parks, but there is no objective basis for this value—it is a human decision about the value of experiencing natural beauty. But the end result is a cost-benefit analysis that looks objective, because we've piled a lot of accounting on top of that recreational value.

There is another obscure-sounding but utterly crucial example of burying the value, and it's pervasive throughout all cost-benefit calculations: *the discount rate*. Whenever we are doing a cost-benefit analysis of present costs versus future gains, even if every other number is entirely objective, there is inevitably one crucial number that is not. The discount rate is how much we discount the value of *future goods* compared with *present goods*. Let's say we need to invest a thousand dollars today in a chocolate factory, and we will get a return of eleven hundred dollars next year. Should we do it? Is it worth losing control of that amount of money now to get that slight profit later? The answer is not an objective matter, because it depends on how we set the discount rate—how we decide to value future goods compared with present goods. If we set the discount rate at 0 percent, then we're saying we think a future potential dollar is precisely as valuable as a dollar in the hand right now. In that case, we should certainly invest in the chocolate factory. If we set the discount rate low—at 5 percent per year, say, then we should also still do it.

But if we set the discount rate high—let's say 20 percent per year—then we should not make that investment. By that discount rate, $1,100 in next-year dollars is worth only $880 in present-year dollars. Most economists think setting a discount rate of 0 percent is ludicrous, because it licenses almost any investment with even marginal gains—since marginal future gains, multiplied over the massive timespan of the whole future, will outweigh almost any present cost. We need some kind of discount rate, but there is no consensus on what the right discount rate actually is.

Every cost-benefit analysis must set some discount rate, but there is no objectively correct rate. It is a value decision about how much future goods are worth compared with present goods—a decision about the relative value of the future. But it is hidden, so the cost-benefit analysis appears cleansed of subjectivity—an immaculate conception of pure math.

We're hiding from the complexity of our value decisions by following an apparently objective metric. But really, we are deferring to the values hidden within that metric.

On to the second kind of objectivity laundering; let's call it the *objectivity bait and switch*. This is when we take a complex, multidimensional quality, often with a lot of freedom of choice, and substitute a simpler quality that is genuinely more objective. If you are convinced that the simpler measure captures everything important, then the whole process will seem entirely objective. But the value decision is hidden almost out of sight—in the initial decision of which simple quality to sub in.

It can be easy to miss the bait and switch, so let's start with a clear case. I was once talking about this stuff in one of my big lecture introductory classes, and a student from the back row—a dude who definitely worked out—raised his hand and said, "But what if we're just measuring something that's already totally objective? Like if you want to get fit and healthy, weight loss is objectively measurable." And the obvious reply is: "When did we start thinking that health meant weight loss?" And when I said that, the class kind of rippled with nervous shock and laughter. What some of them said after was: It was so obvious once somebody pointed it out to them, but they were shocked that they'd needed it pointed out. Because, obviously, health isn't weight loss, but it's easy, somehow, to get sucked into thinking that way.

The crucial point here is that the value-laden decision isn't hiding in the way we measure weight. Weight really is just a simple, objective measurement. The value-laden decision is one step back, in the decision to use weight as a stand-in for health in the first place.

OK, now for a much more complicated example. Let's look at a hundred-point rating scale that dominates the modern wine world. Wine scorers have tried to make the process as objective as possible. This means a few things. For one, professional wine tasters usually spit out the wine—to avoid getting drunk, so that they can render an objective judgment.

For another thing, professional wine tasters, when officially scoring a wine, have it *without food*. This is because having wine with food would create too much variability in the judging process. A wine that might be incredible with a roast chicken might be only decent with fried chicken, and might disappear next to a peppery steak. So, to create a more stable, objective judgment, the official wine-judging process excludes food.

But for many people, much of the joy of wine is how it reacts with food. Wine can change and shift with every sip, reacting dynamically to different bites. And there are different relationships it can have with food; it can be a sharp contrast or a resonant match. But this level of variability would create incredible difficulties for an official scoring process. How could you score a "best" wine if one is incredibly good with tomato sauces but kind of listless with anything else, and then rank it against a wine that is moderately good with every food? So we create an artificially stabilized context; we cut out the variability. We score wine without food. But if we follow that wine-scoring system, we have implicitly adopted a value choice: that objective scoring is more important than the relationship between wine and food.

And that value choice feeds back into the wine itself. The wine-scoring system is transforming the winemaking world, pushing it toward "fruit bombs"—wines with big, bombastic flavors. And since high scores mean better sales, many winemakers have started making more and more fruit bombs. But those wines, though they perform better in the sterile, controlled environment of the wine-scoring process, usually don't react in interesting ways to food. If you grow up in the era of these new fruit bombs, though, you might never know that wine could react so powerfully and variably to food. The wine-scoring system has written itself into reality by transforming the wine-growing industry, blotting out dynamic wines and filling the world with stable and unreactive ones.

OK, now for the most complicated example. Let's get back to the concept of "healthy." We already saw one mistake—substituting weight loss for health. This is a common mistake, but relatively easy to see. But health is also subject to more sophisticated forms of the objectivity bait and switch.

The value decision is hidden almost out of sight—in the initial decision of which simple quality to sub in.

The philosopher Elizabeth Barnes suggests that our metrics and measures for health often end up obscuring our free choice under the veil of science. People act as if there is a single objective quality—health—and a single objective measure for it. If that were true, then health would be something we could measure scientifically—and then establish a singular and objective list of best practices. But this is a mistake, she says—a misunderstanding of what health is. There will, according to Barnes, never be a single clear metric for health, nor will there be a one-size-fits-all set of best practices for all people to pursue health. This is because health is not a static thing, stable across all people. It is a concept that contains, hidden within it, some freedom to choose. What health is depends on a complex set of decisions about what we value about our bodies and capabilities. And different people will make those decisions in different ways.

First, a quick primer from the philosophy of language. Some words have a stable meaning. The meanings of the words *apple* and *four* will usually stay the same in different contexts. But other words have a meaning that varies with context, like *tall*. Am I tall? That depends. Am I tall for a human being? I'm a bit taller than the average human. Am I tall for a professional basketball player? Absolutely not. Such terms change meaning depending on a comparison class. Terms like *fit* and *healthy* can also work like this. Am I fit? That depends. Compared with the average academic philosopher, I'm fit as hell. Compared with the average rock climber, I'm a weak slob.

The meaning of *health* has a very special form of contextual variability, says Barnes. The meaning changes from person to person depending on each person's *interests*. I ask my doctor: Is my knee healthy? Does my current exercise routine support a healthy knee? But there is no single objective answer to that question because the meaning of *health* fluctuates, depending on my interests and goals. There are limits to that flux: No matter your interests, being run over by a bulldozer is unhealthy for your knees. But within those limits, the exact meaning of "health" depends on your particular interests.

Suppose Samantha is an Olympic sprinter. She is interested in whether she will be able to stand up to the rigors of competition and have a chance at a gold medal two years from now. If she has to trade off some late-life functionality and pain for short-term success, she'll do it. Health for her involves short-term high-end athletic functionality. I, on the other hand, am a casual rock climber. I climb to relax, for mental health, and because I love the beauty of the movement. And I run mostly as cross-training, to balance out the abuse that climbing puts on certain joints. I am most interested in whether my knees can keep me climbing well into my sixties, which will include regular ten-foot falls onto a gymnastic pad.

If I have to accept some pain in order to continue climbing as long as possible, I will, and I am willing to sacrifice short-term optimal performance for long-term functionality. Health for me involves balancing longevity with moderate athletic functionality, while accepting some degree of pain. Or you might be interested in how long you can continue to walk pain-free, as late in life as possible. Health for you involves long-term functionality, but weighting pain reduction over athletic performance.

The notion of health is essentially fuzzy, says Barnes. It's an approximation—the vague center of gravity of a cluster of different people's interests. There are a lot of things that the sprinter, the climber, and the walker will have in common, and there are some broad generalities that will generally promote knee health. Losing all your knee ligaments is bad news for pretty much any human, whatever their interests, and drinking water is generally a good support for any human functioning. So the rough approximation is good enough for some uses. But that kind of universal prescription doesn't work in every case. This is why there are no true best practices for health—because the right thing to do depends on each person's different interests. The notion of health is messy, says Barnes, because human interests are messy.

But if we take a simple, genuinely objective measure and use it as our meaning of *health*, we can hide from the need to make choices. And we have some very promising candidates for the bait and switch. There are some qualities that are affiliated with many conceptions of health that are easy to measure mechanically, such as longer lifespan and lower heart attack rate. They are excellent candidates for standardized measurement. And certainly they are decent as rough approximations for our shared interests. But they are incomplete.

Let's say we set the meaning of *health* in terms of some measurable qualities: longer lifespan and lower incidence of measurable disease and injury. Even this is a value decision; it sets certain interests above others. It sets an interest in increasing the number of years and ignores the interest in the quality of those years. It sets an interest in avoiding disease over an interest in developing complex physical skills. It will tell me to run on a treadmill instead of going trail running on difficult terrain, because trail running is riskier and increases the chances of significant injury—and this particular notion of health values avoiding injury over the development of complex motor skills.

If we take a simple, genuinely objective measure and use it as our meaning of health, we can hide from the need to make choices.

I'm not arguing that we can't measure anything objectively or that science doesn't work. Some kinds of things can be measured precisely, because their nature doesn't vary with interests. We can measure, with perfect objectivity and no information loss, inches, pounds, and calories. But concepts like "health" are fundamentally unlike "calories." A calorie is independent of your values and interests. Health is dynamic and variable, because it is essentially entangled with your interests. To be healthy is to have your body and mind functioning as you wish. This involves a choice about which aspects of your functioning you care more and less about. But we can launder the concept of "health" and make it look entirely objective. And if we accept that laundered "health," we will be letting somebody else make value decisions about how we should take care of our own bodies and minds.

Thinking about real and fake objectivity also helps us to answer one of our big questions: When should we trust metrics, and when shouldn't we? When should we be worried about large-scale data collection, and when should we be fine with it?

Very early in the *Nicomachean Ethics*, Aristotle warns that we should not demand too much precision when it's inappropriate. Some areas, like engineering and science, allow for great precision. But you can't expect the same degree of precision in politics and ethics. It is, says Aristotle, the mark of the educated person to seek the appropriate amount of precision for each subject. And you can't expect precision in politics and ethics, because what we should do in these depends significantly on human autonomy and choice.

To sum it up for the data and policy wonks in the audience, large-scale data collection efforts will likely be most useful when they target those qualities that:

1. are highly *invariant* and *stable* between contexts

2. have accessible, mechanical methods for being counted

3. do not involve a value decision

One of the reasons large-scale data collection works so well for figuring out antibiotics is that antibiotics work in a relatively context-invariant way. They are effective on most human bodies, regardless of the place, psychology, or culture of the person. And the positive results of antibiotics are easier to measure, because we can measure the disappearance of bacteria mechanically.

Large-scale data collection efforts will more likely miss the mark for qualities that:

1. are highly *variant* between contexts and dynamic

2. require a nonmechanical exercise of skill, sensitivity, or expertise in order to be counted

3. involve a value decision

Metrics will not capture the goodness of wine because many important qualities are highly dynamic and variant between different dynamic contexts. They will not capture health because health involves a value decision at its core. **We can get rough approximations that are helpful for large-scale population work, but we should not take them to be complete or decisive for individuals.** 🔒

Info

Ready to dig deeper into the book?
Buy a copy of <u>The Score.</u>

Want copies for your organization or for an event?
We can help: customerservice@porchlightbooks.com

800-236-7323

ABOUT THE AUTHOR

C. Thi Nguyen is associate professor of philosophy at the University of Utah, and a specialist in the philosophy of games, the philosophy of technology, and the theory of value. A former food writer for the *Los Angeles Times*, Nguyen is active in public philosophy, writing for *The New York Times*, *The Washington Post*, *New Statesman*, and elsewhere.

SHARE THIS
Pass along a copy of this manifesto to others.

SUBSCRIBE
Sign up for e-news to learn when our latest manifestos are available.

Porchlight

Curated and edited by the people of Porchlight, ChangeThis is a vehicle for big ideas to spread.
Keep up with the latest book releases and ideas at porchlightbooks.com.