

**The Limits of the Numerical:
Metrics and the Humanities in Higher Education
Christopher Newfield
University of California, Santa Barbara
DRAFT December 2015**

In his classic book, *The Audit Society* (1997), Michael Power noted that metrics-based governance is inadequate, and then suggested that metrics-governance would soon run its course. This is the only false note I've found in that book, which elsewhere includes Power explaining why this optimism would not pan out. The reason why is that the failure of audit to achieve its stated goals—like purging waste or lowering costs—is always followed by renewed applications of audit to achieve these goals. Failure is a feature of audit, not a bug. In the context of audit's boundless powers of self-renewal, we need to answer these questions: What do we do when audit is never finished? What do we do when there is no non-audit normal for things to go back to?

In the US, where audit culture is weaker, the political management of universities has fallen to techniques like the long-term culture war against universities and its foundation of professional expertise. The culture war first targeted ethnic studies and gender studies for generating political correctness rather than validated knowledge, but since 2000 the targets have included atmospheric scientists that work on climate change, medical researchers that use stem cells in their laboratories, among many other examples—and research on the effects of tobacco or of lead as a fuel additive go back much further than that. While the U.S. culture war expressed hostility to identity-based social movements and the war on science expressed hostility to critiques of deregulated business, the UK's political management of universities has used audit in the name of democracy, deploying the free market language of quality assurance and consumer choice. The Tories, in contrast to the US Republicans, tie audit to democratic values in ways that can disarm the critique of audit culture and its metrics. A further complication is that all major political parties in the

English-speaking world favor audit. The program to close U.S. public schools on the basis of student test scores was pushed by the Obama administration, and the proposal to replace the U.K.'s Research Excellence Framework's peer-review with bibliometrics came from New Labour's Gordon Brown. In short, critics of audit cannot dismiss or ignore its democratic claims.

I would eventually like academics to have alternative forms of collaborative evaluation that embed quantitative indicators in public life, professional practices and intellectual needs. But there is the question of why we should go to the trouble? Would alt-metrics be better? What exactly do current metrics lose in the way of knowledge, thought, institutional relations, and the like?

These are empirical questions that existing research has only partially answered. I am part of a three-year multi-national research project on this topic and we will have more to report a few years hence. This is a kind of gateway essay. In it I'll invoke one driver behind metrics in higher education, discuss leading general criticisms of metrics, and end by talking about research questions whose answers might get to a post-metrical system of humanities-friendly evaluation.

1. A Motive for Metrics

Numerical indicators are particularly suited to *economization*, which has been the master strategy under austerity, which is in turn reinforced by economic confusion and weak recovery in the West. Over the past forty years we've squeezed a lot of new wealth out of digital technology, and out of the financial sector, but this wealth has been unequally shared—in most accounts not shared much at all. Few policymakers think the US or the UK economies are on the verge of generating the steady growth in productivity and broad prosperity that Western societies saw after World War II. University policy is being made in a context of permanent policy anxiety about the economic future.

This of course means that universities around the world are having to adapt to contradictory demands – to be better *and* be cheaper. Universities are increasingly being held responsible for the state of their nation's knowledge-based economy, for the workforce readiness of the youth

population, for the technologies that are to found new industries, and for social cohesion and political capability. As university performance has become a major economic and political issue, a range of quantitative metrics has developed to assess and govern that performance.

This trend is particularly well-studied in the United Kingdom, Australia, New Zealand, and the European Union, where New Public Management and productivity audits are established university practices and have large literatures in various languages.¹ The U.S. also has a policy tradition that goes back to “reinventing government” in the early Clinton administration, and its own practices of scientific research assessment.² What we could call academic knowledge performance has become the domain of quantitative procedures that are linked together in a complex assessment culture. Quantitative assessments continue to multiply.

Although much about the use of numerical indicators is not new, there is much that is. Its comparative *scope* is new: it now furnishes the basic picture of a very heterogeneous international higher education sector, in part by reducing or eliminating contextual variety. The numerical’s *priority to qualitative* considerations is new, and has resurged in spite of decades of criticism of the validity of traditional quantitative metrics in achievement testing and other areas.³ The numerical has gained new momentum through the worldwide interest in “big data,” where advanced data processing techniques appear to offer a simplicity, convenience and *objectivity* not found in qualitative comparisons. The numerical has the ability to establish *benchmarks*, which must favor processes and goals that are more measurable over those that are complicated or inseparable from specific contexts.⁴ (Ordorika and Lloyd 2013). In a broad cultural shift, quantitative benchmarks, rankings, and indexes have also acquired a *universality and inevitability* that give them quasi-automatic authority over educational governance. One particularly interesting feature is that they shape resource decisions and public policy even when their validity is disputed.⁵

It doesn’t overstate the case to say that in higher education’s current global phase, quantification has expanded from being used under special circumstances for specific insights about complex situations to being available under all circumstances as a privileged, decisive, and objective mode of comparison and assessment. The *qualitative* analysis of learning,

research, or institutional resource allocation has come to seem partial, specific and subjective by comparison to quantitative analysis, which most decision makers now regard as more general and objective.⁶ Academic research and institutional decision-making require comparison of dissimilar elements across varying contexts, and higher education now generally assumes that the numerical is the best adjudicator of comparisons across multiple differences.

So it is a matter of some urgency to clarify the relation between strengths and *costs of the numerical* as universities try to adapt to conditions that are changing in different ways across the globe—and to propose better methods of evaluating and guiding these adaptations where we find fault with the current varieties. This is particularly true given audit's fit with austerity, which blocks investment in higher education upgrades across the entire spectrum of institutions.

2. Metrics as Audit Democracy

One critique of audit is that it exposes professionals to external management. The professions have long been groups that seek to govern themselves, and mostly with standards and evidence that are incomprehensible to outsiders. Laypersons don't understand the symptoms of malpractice in a cardiac surgery unit, or whether a 1960s engineering error made the World Trade Center towers too flammable when hit by passenger jets. Nor do their elected representatives. But each can compare the fatality rates of their local cardiac unit to one further away, or a grading system for the flammability of construction materials. Hence metrics like fatality rates are in common use in assessing health care quality.

For the professional, the problem is not so much the quantified result, reductive and misleading though that may be, as it is the managerial apparatus that accompanies it. The apparatus comes to micromanage the behaviors surrounding the activity being measured, whether it is cardiac surgery or teaching 17th century English poetry. One way of naming this phenomenon is as bureaucratic management. It not a new problem by any means—Weber himself worried about the loss of individual authority. But the problem seems to be getting worse by the week, and one person who has recently discussed it is the anthropologist David Graeber, in his book *The Republic of Rules*. Here's one thing he says:

We are witnessing a direct clash between two different forms of materialized utopianism: on the one hand, an anti-authoritarianism that, in its emphasis on creative synthesis and improvisation, sees freedom basically in terms of play, and on the other, a tacit republicanism that sees freedom ultimately as the ability to reduce all forms of power to a set of clear and transparent rules.

For the last two hundred years, in Europe and North America—and increasingly, elsewhere—that latter, bureaucratized notion of freedom has tended to hold sway. New institutional arrangements that operate by rules so strict and predictable they essentially disappear . . . They tend to be put forward as platforms for human freedom that emerge from the very technical contingencies of running efficient structures of power.⁷

Graeber doesn't simply contrast creative freedom and managerial tyranny, but points out that management is a venerable mode of freedom—freedom as due process and equality under rules. This gives audit culture a justification, even if it does crush “creative synthesis and improvisation.”

Audit culture in fact defines metrics as the heart of a democracy and equity movement that subjects various professional priesthoods to the light to day. To deal with audit, we need to see clearly that it claims three progressive features.

1. Social Integration: in the last few centuries, standardization has become important--more important than accuracy, in part because it enables communication across the differences that became constitutive of modern societies.⁸
2. Political Democratization: Metrics allow citizen access to the outputs of expert elites through apparently common and transparent rules. There are intellectual as well as ethical benefits: “This preference for rules over unconstrained judgment is common both within science and in the larger political arena. The free exercise of judgment invites suspicion of arbitrariness or bias.”⁹ Those advantaged by special access to knowledge and power must, through metrics, report to regular folks.

3. Economic Equity: Alan Williams is one of the originators of the QALY, or Qualitatively Adjusted Life Year, which sets a numerical cost-benefit threshold for different clinical treatments, set in terms of the value of a relatively healthy year of a patient's life. For instance, if a treatment costs less than £20,000 per QALY, it is more-or-less automatically approved. If it costs more, various kinds of protocols and evaluations kick in. Though there are serious issues with the methodology, Williams' *motive* appears to have been equity: before a clinician does everything possible to save one person's life, it is good to know how that expense will damage efforts to save other people's lives who are not consulted. The QALY is part of a numerical system that claims to place any one person's treatment costs, no matter how wealthy, lucky, or well-placed, in the context of its impact on society's treatment resources. Williams thus doesn't mind the term healthcare "rationing," noting that "the term "rationing" [was] to me (having been a wartime schoolboy) benignly synonymous with distributing scarce resources equitably."¹⁰

We can perhaps now better appreciate the power of the Conservative government's use of democracy language in the Browne report that revolutionized university funding in the UK, and in the Green Paper that seeks to apply a national audit system to teaching. Here's some Green Paper text:

18. We want to see greater assurance for students and employers that the class of degree awarded by higher education providers is consistent across the sector, provides a meaningful differentiation between undergraduates that recognises variation in levels of attainment, and that the value of a degree awarded remains the same over time.

21. . . .The new Teaching Excellence Framework (TEF) aims to recognise and reward high quality teaching. Prospective students will be able to use the TEF results to help inform their decisions about which institution to attend, and employers can consider it in their recruitment, rather than relying on the more imperfect proxies for teaching quality such as research success. The TEF will increase students' understanding of what they are getting for their money and

improve the value they derive from their investment, protecting the interest of the taxpayer who supports the system through provision

A.1.2 More needs to be done to ensure that providers offering the highest quality courses are recognised and that teaching is valued as much as research. Students expect better value for money; employers need access to a pipeline of graduates with the skills they need; and the taxpayer needs to see a broad range of economic and social benefits generated by the public investment in our higher education system. That is why we are developing a new Teaching Excellence Framework.¹¹

The Green Paper uses the language of providing access, boosting social mobility, including the disadvantaged, offering student choice, fulfilling our potential—it's Matthew Arnold as a school inspector handcuffed to the California ed-tech guru Sebastian Thrun. At the time, I battled Thrun's MOOC vision tirelessly—there are all sorts of technical and financial problems with it.¹² There are also psychological problems: I am quite interested in what I have started to call *audit suffering* -- I think it induces Melanie Klein's "depressive position" in academic staff. And yet the coercion inflicted on professionals—which can reduce creativity, damage productivity, and destroy the pleasures of work—has a democratic rationale. This keeps everyone thinking well, some kind of audit is a good thing in moderate doses. In my experience, nearly all of the British professoriate thinks this too.

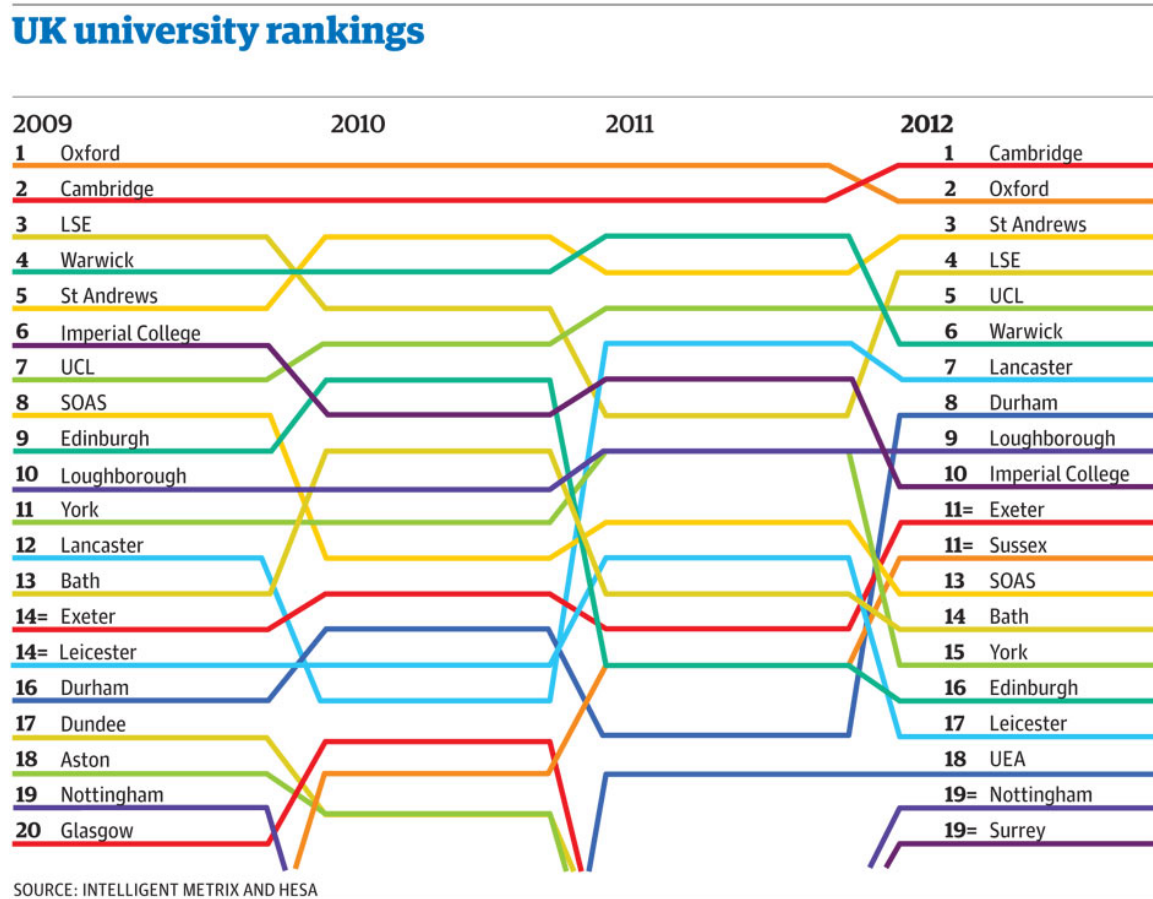
I do think what we might call audit suffering is real. But it seems to exist for the sake of democratic accountability. In the case of this kind of trade-off, the default is to continue to use audit and numerical indicators with caution and qualification. As the authors of the "Leiden Manifesto for research metrics" put it, "The best decisions are taken by combining robust statistics with sensitivity to the aim and nature of the research that is evaluated. Both quantitative and qualitative evidence are needed; each is objective in its own way."¹³ Is this a good stopping point—using numerical metrics in conjunction with a judicious interpretative context? I don't think so. We need to understand not only the strength of numerical indicators, but effects that seem to be inseparable from their use.

3. Critiquing Metrics: Epistemological Gaps

A common criticism of metrics is that they are inherently inaccurate because they never quite measure the thing they say they measure. More specifically, they measure proxies for a thing rather than the thing itself. The relation between the proxy and the thing is imperfect and sometimes inverted. Gaps between signifier, signified, and referent has been a core issue in literary theory for forty years, and descends from the semiotics of Peirce in the 19th century. It is a well-understood problem that numerical indicators take to a new levels. I'll use university rankings as a straightforward example.

Here's an example of a kind of proxy effect.

Figure 1: UK University Rankings (Guardian University Guide).¹⁴



It's fairly interesting to watch a particular university go up and down. There's the drama of the transition from 2011 to 2012 where Oxford and Cambridge swapped positions 1 and 2. Is Oxford really number 2 rather than number 1? Is Cambridge really number 1 rather than number 2? Will

they swap again? Not so far, but you never know. Then there's the London School of Economics, which dropped and then bounced back. Imperial College of London and University College of London traded places. Nottingham plunged into the drink—and then climbed out to rest exactly where it was before. Dundee and Aston disappeared, perhaps forever. The House of Lancaster defeated the House of York. And so on. This kind of variability reflects some mixture of market reputation and more or less successful efforts to influence the indicator, rather than a real change in the underlying behaviors of these institutions.

For example, universities that are highly ranked year after year might be assumed to offer excellent student engagement. All ranking systems try to get at teaching quality. US News's factor mix now looks like this.

“Graduation and freshmen retention rates count for the largest share of the 100-point ranking, 27.5 percent, with extra points for schools whose graduation rates exceed statistical expectations based on socioeconomic mix and other factors.”¹⁵ The mix is reasonable.

Figure 2: US News Ranking Factors (University of Tulsa graphic)¹⁶



But all the student-related material—“Quality,” “Students” and “Graduation” rests on proxies for learning—class size, graduation rate, and freshman retention. These are important, but they don’t tell you directly what anyone is learning, or whether anyone is engaged. A study this year by two Indiana University scholars found no correlation between institutional ranking and student engagement. It found an *inverse* correlation between ranking and student-faculty interaction, meaning the more highly ranked the university you attended, the less likely you were to know your professor.¹⁷ And of course engagement is itself only a proxy for learning, where studies suggest another non-correlation with ranking.

A central reason that rankings are not reliable is that institutions focus not on improving the thing the metric supposedly measures but on raising the metric itself. A barely second-tier private university in Boston called Northeastern took the bull by the horns:

Starting in the 1990s, Northeastern University set out to reverse-engineer US News' formula. . . Then they did everything they could to climb the rankings. They lowered the enrollment cap on classes to 19, because 20 students is no longer seen as a small class in US News' eyes. They began using the Common Application, which makes it easier for students to apply — at a cost of \$75 per application, of course — so that they could turn more students away. When that wasn't enough, they spent liberally on recruiting in the US and abroad, hoping to become ever more selective.

Admitted students with so-so high school grades and SAT scores were encouraged to study abroad for their first semester in a newly created program, because students who start in the spring don't count for US News purposes. Meanwhile, the university's former president went out of his way to meet leaders of highly ranked colleges and even lobbied US News itself to tweak its formula to benefit Northeastern.

It worked. Northeastern rose from 162nd in 1996 to 49th [in 2013].¹⁸

This story displays two features of the use of metrics in higher education. The first is that the subjects of an indicator can begin to shape their behavior to improve the indicator. This corrupts the indicator, as it no longer refers, even as an imperfect signifier, to behavior that is independent of the indicator. The second is that thought and resources are invested in influencing the indicator rather than in fixing the thing the indicator measures. It appears that Northeastern capped seminar size just below U.S. News' arbitrary quality threshold and spent money on facilities that took weak first-year students off the books. Neither of these activities addresses the thing supposedly measured, which is student learning.

There's a third important issue here. Improved rankings are seen by governing boards and related influential figures as an index of entrepreneurial spirit. Rankings have an authority derived in part from their power of social and cultural signaling of a desirable quality—in English-speaking cultures, the entrepreneurial self has become an unqualified good, almost a state of grace, a condition of personal superiority to others. The contrapositive is also true: a senior manager that does not seek improved rankings is inferior to those who do. This is a

distinctive cultural belief that is very difficult to criticize or even examine systematically.

Taking these three issues together, a critique of the numerical in university rankings might start by suggesting that the politics (including the managerial self-interest) of metrics overrides unresolved methodological and philosophical problems. We can't say for sure what this has done to research or instructional quality.

4. Metrics Against Social Equality

Most policymakers and administrators believe metrics are necessary and inevitable, and the main response to criticism of college rankings over the past thirty years has been the multiplication of rankings—and also the multiplication of components that can be adjusted within any given ranking. All rankings have tried to move from subjective to objective factors by shifting from reputational surveys toward bibliometrics for research, and toward quantitative student outcomes for teaching.

Everyone also tries to customize their particular rankings. *Forbes* magazine favors student satisfaction and income return on investment. Its rankings weigh "Student Satisfaction (22.5%) [survey data], Post-graduate success (37.5%) [personal income], Student Debt (17.5%), Graduation Rate (11.25%) and Nationally Competitive Awards (11.25%)." ¹⁹ The good news is that rankings are now self-conscious about how shifts in methodologies change the results. They are self-conscious that these rankings do not reflect an objective, independent reality but instead measure factors the rankers have decided are important.

This point has often been made. Rankings gurus conscientiously adjust their formulae. What is the result?

Here's one picture – theunipod's averaging of everybody else's rankings.

Figure 3: theunipod Ranking of University Rankings 2013--US universities²⁰

Rank	University	THE *	ARWU *	QS *	US News
1	Harvard University	2	1	2	2
2	Stanford University	3	2	3	5
3	Massachusetts Institute of Technology	4	4	1	7
4	Princeton University	5	6	6	1
5	California Institute of Technology	1	5	6	10
6	Yale University	8	9	4	3
7	University of Chicago	7	8	5	5
8	Columbia University	10	7	9	4
9	University of Pennsylvania	12	13	8	7
10	University of California, Berkeley	6	3	14	20
11	Johns Hopkins University	11	15	11	12
12	Cornell University	15	11	10	16
13	Duke University	13	23	13	7
14	University of California, Los Angeles	9	10	17	23
15	Northwestern University	16	22	15	12
16	University of Michigan	14	18	12	28

The difference made by various weightings is no difference at all. The same universities are always at the top no matter what you do to the standard factors.²¹ At least ten of them would have been in the top of the same list in 1899. Research has confirmed that standardized test scores do correlate with family income and do *not* correlate with future university performance. There's a similar phenomenon here. The top schools are mostly the old schools, the rich schools, the private schools, and also the schools that were already the top schools, which are also the schools that can *buy* top-ranked candidates in the annual auction for high-testing students that keep them top schools. Rankings measure, to a large extent, *incumbency advantage*.

Rankings also help people mistake the tip of the iceberg for the iceberg itself. The top 20, or top 200 universities are not an accurate synecdoche

for the other 4000 or so. The media analyst Clay Shirky made the point in an effectively brutal way.

Imagine picking a thousand students at random from among our institutions of higher education. Now imagine unpicking everyone at one of *US News'* Top 100 [liberal arts colleges](#) or [universities](#). [N.B. He must mean 200 schools total.] You'd expel anyone from the Ivy League, Stanford, MIT. Anyone from Emory or Rice. Anyone from Vanderbilt, Clemson, Drexel. Anyone from the famously good state schools—UMass, Virginia, the California universities. After ejecting those students from your group, how many of the original thousand would be left?

About 900.

The total enrollment on those two lists—which includes almost every good college you've ever heard of and some you haven't (Wabash; SUNY College of Forestry)—only accounts for a tenth or so of the 18 million or so students enrolled this year at one of our thousands of institutions of higher education.

If you want to know what college is actually like in this country, forget Swarthmore, with 1500 students. Think Houston Community College, with 63,000. Think rolling admissions. Think commuter school. Think older. Think poorer. Think child-rearing, part-time, night class. Think 50% dropout rates. Think two-year degree. (Except don't call it that, because most graduates take longer than two years to complete it. If they complete it.)

If you want to know what college is actually like in this country, skip Google Images, and scroll through the (still heartbreaking) [We Are The 99 Percent](#) Tumblr, looking for the keywords "student loan."²²

In the US, rankings help people overestimate the quality of the US system as a whole. They also class at least 90% of college enrollments as enrollments at unranked colleges. In the UK, they designate the majority of enrollments as inferior to the "world-class" portions of the UK system.

Let's assume for the moment that these points are true. Just as mental aptitude tests turned out to reflect social rather than biological inequalities,

rankings reflect incumbency advantages rather than internal institutional superiorities. And then, rankings overstate the quality of the overall system while relegating the majority of students to status inferiority.

These are examples of a core criticism of metrics, which is that they normalize or naturalize social inequalities. There is then a fork between milder and stronger versions that critics of metrics can draw. The milder claim—normalization-- is that metrics make everyone think that the university system has always been stratified like this. This then distracts policymakers from addressing real problems starting with funding inequalities and including racial inequality and climate. This is an important issue: it is encapsulated in a joke in which a state legislator supposedly says to public university president, “We used to send you money. Now we send you rules.” Audit’s political claim is that universities have always had plenty of money and now they can do without funding increases by eliminating waste. We have two generations of politicians refusing to make correct public investments with inefficiency as an alibi and audit as both diagnostic and cure.

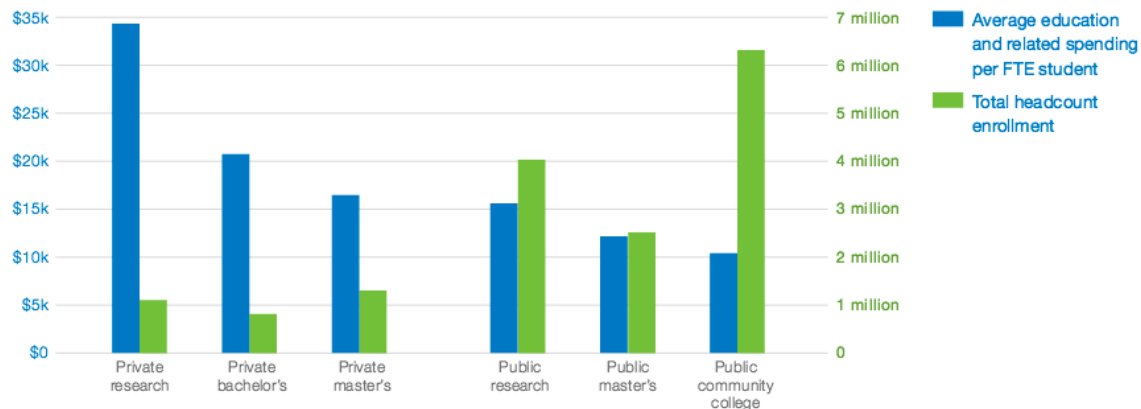
The stronger claim is that rankings *naturalize* funding inequalities to the extent that rankings claim to measure the intrinsic merit of students, faculty, and the overall institution. If one institution has less money per student than another, it reflects their lesser quality. Were you a good enough student, you’d be able to get into a better and generally richer school; if you’re at a poorer school, that reflects your poorer achievement.

In the US, these wealth gaps are dramatic. Here’s one illustration.

Figure 4: Average Educational Spending per student vs. Enrollment, 2008²³

Institutions serving the most students spend the least amount on their education

Enrollment vs. spending, AY2008 (in 2008 dollars)



Source: Delta Cost Project IPEDS database, 1987–2008, spending data from 11-year matched set, enrollment data from unmatched set.

These are averages: actual cases are more extreme: community college students, likely to be socioeconomically disadvantaged, would receive one-tenth the per-student resources of a private university. A 2013 study at Georgetown University found that the non-selective students that receive the majority of new African American and Latino students in the past two decades spend 1/5th of the per student expenditure of the 468 “selective” universities that received the great majority of new white students. The US Supreme Court is again reviewing a race-based affirmative action program, and Justice Scalia suggested during argument that Black students might be better off at “slower” schools. What he *also* means in practice is that they’d be better off at poorer schools. We know that less money for individual advising, supported peer programs, academic tutoring, and so on lead to lower graduation rates.

So: the strong version of the claim that ranking naturalizes inequalities suggest that rankings help cause social (and racial) inequality. Rankings *cause* the Matthew Effect, in which the rich get richer, by claiming to measure performance superiorities inherent in particular institutions. Numerical indicators, in other words, supports rank meritocracy—the idea that whenever you increase merit through education you naturally find a ranked hierarchy of quality. To repeat, this argument is that metrics are defective to the extent that their operation doesn’t simply measure inequality, but naturalizes it, justifies it, and produces it.

This point needs empirical research. But the strong version of the inequality critique posits a causal effect between the audit explosion of the last three to four decades and inequality boom in the US and the UK: Metrics in a range of domains, and particularly in that famous social escalator higher education, has either distracted us from remedies to (the mild claim), or actively naturalized and supported, the Anglophone world's growth of inequality in wealth and income—and in educational attainment at different incomes.

Here's a summary table of the argument so far.

Figure 5: Problems with Metrics

1. When behavior directly addresses the indicator designed to measure it, it ceases to measure the behavior itself
2. Resources are more likely to focus on changing the proxy than on changing the thing the proxy is to measure (e.g., class size, not learning)
3. Audit signals entrepreneurial spirit to authorities, artificially increasing its authority
4. Audit generates accountability by inducing suffering that reduces “creative synthesis and improvisation”
5. Audit's democratic rhetoric is undermined by either enabling or causing anti-egalitarian effects

I don't list these problems as final conclusions or rules. I intend these issues to point towards research questions that need empirical as well as conceptual investigation. A billion-dollar industry has addressed many technical questions but ignored others, particularly philosophical and socio-cultural questions. Is audit suffering confined to intellectuals, teachers, physicians, and other professionals, or has it spread throughout society? Does it not only justify workplace surveillance and the kind of but also cause stratification and scarcity? Can we make knowledge economies transparent and democratic without the negative effects that current numerical indicators bring us?

5. Coda: Metrics and the Global University

Rather than concluding with a bit more about how our selected metrics work and what their impacts are, I'll say something about why it matters. My experience is that audit culture is helping to cloud the imaginations of the humanities disciplines unnecessarily, to place humanities knowledge on the margins of universities and their wider societies, and that it is doing these things at an unfortunate historical time.

We are living through a paradox in current global university development. On the one hand, universities are being asked to adapt to a diverse range of local, national, and regional needs in a world where knowledge economies are following different pathways of advancement and evolving their own rules and norms. It is increasingly clear that economic development can no longer be implemented by exporting and imposing a single model devised in the West (or anywhere else). As the anthropologist Harriett Moore has observed,

if we focus the future on prosperity, on broader understandings of human flourishing, rather than on economic development based on growth, we will have to accept that the prosperous societies of the future will be very diverse. In such contexts as these, learning will not be about the transferability of whole models with known outcomes, but rather about incomplete learning, experimentation and collaboration. (Moore, 2015)

And yet, at this historic juncture, numerical indicators create uniform standards to which universities must norm themselves in order to succeed with global metrics. Universities may be making themselves *less* functional in their home “conditions of production” by improving functionality in areas subject to quantitative measurement. The shrinking of the humanities is a symptom of a declining model of development, one to which too many university leaders mistakenly cling.

In reality, universities cannot serve their societies by adopting one model of education and management. They must instead be sources of multiple design pathways, diverse methodologies, and even multiple rationalities. To provide these needed varieties of thought and practice, universities will need to take multiple forms, and have the capacity for “experimental

collaboration” with their regions and with parts of the population that they have not traditionally served. We, professors of physics and literature alike, must help our university managers reject metrics that suppress diverse knowledges and knowledge practices as a simple matter of sustainability.

Universities going forward will be *adaptive* to changing local, regional, and national needs. They will generate *diverse* knowledges, practices, and skills in their graduates. They will *individualize* their students by using technological and human systems to reverse the massification and anonymization that is known to impair learning. Universities will *reduce status and resource stratification*. Whatever the past of metrics culture, the future belongs to knowledge diversity and experimentation. And of course, as literature professors, we *specialize* in that.

NOTES

¹ Bowker, Geoffrey C., and Susan Leigh Star. 2000. *Sorting Things Out: Classification and Its Consequences*. Cambridge, Mass.: The MIT Press; Bruno, Isabelle, and Emmanuel Didier. 2013. *Benchmarking*. Paris: Zones; Lampland, Martha, and Susan Leigh Star, eds. *Standards and Their Stories: How Quantifying, Classifying, and Formalizing Practices Shape Everyday Life*. 1 edition. Ithaca: Cornell University Press, 2008; Power, Michael. 1997. *The Audit Society: Rituals of Verification*. Oxford; New York: Oxford University Press; Strathern, Marilyn. 2000. *Audit Cultures: Anthropological Studies in Accountability, Ethics and the Academy*. 1 edition. London New York: Routledge; Vinokur, Annie. 2008. "La loi relative aux libertés et responsabilités des universités : essai de mise en perspective." *la Revue de la Régulation* 2 (Janvier).

² Gore, Al. 1993. *The Gore Report on Reinventing Government: Creating a Government That Works Better and Costs Less*. New York: Three Rivers Press.

³ Lehman, Nicholas, *The Big Test: The Secret History of the American Meritocracy* (New York: Farrar, Straus and Giroux, 1999); Ravitch, Diane. *Reign of Error: The Hoax of the Privatization Movement and the Danger to America's Public Schools*; Kamenetz, Anya. *The Test: Why Our Schools are Obsessed with Standardized Testing—But You Don't Have to Be*. New York: Public Affairs, 2015.

⁴ Ordorika, Imanol and Marion Lloyd. "A Decade of international university rankings: a critical perspective from Latin America." In Marope, P.T. M., P.J Wells and E. Hazelkorn, eds. 2013 *Rankings and Accountability in Higher Education: Uses and Misuses*. New York: UNESCO Publishing. Pp. 209-31.

⁵ Osterloh, Margit and Bruno S. Frey. 2010. "Academic Rankings and Research Governance," Working Paper No 482. Zurich: Institute for Empirical Research in Economics, University of Zurich: 1-29.

⁶ Gates Foundation. *Postsecondary Success Advocacy Priorities*. Seattle: Gates Foundation, 2015.

⁷ Graeber, David. *The Utopia of Rules: On Technology, Stupidity, and the Secret Joys of Bureaucracy* by Graeber, David (2015) Hardcover. Melville House). Pp. 203-04

⁸ Porter, Theodore M. "Making Things Quantitative." *Science in Context* 7:3 (1994): 389-407.

⁹ Ibid.

¹⁰ Williams, Alan. "Discovering the QALY: Or how Rachel Rosser changed my life." MS.

¹¹ UK Department for Business, Innovation & Skills. *Fulfilling Our Potential: Teaching Excellence, Social Mobility and Student Choice*. November 2015. Emphasis added. <http://wonkhe.com/wp-content/uploads/2015/11/2016HEgreenpaper.pdf>

¹² Newfield, Christopher. "Where Are the Savings?" *Inside Higher Ed*, June 24, 2013. <http://www.insidehighered.com/views/2013/06/24/essay-sees-missing-savings-georgia-techs-much-discussed-mooc-based-program>.

¹³ Hicks, Diana, et al. "The Leiden Manifesto for research metrics." *Nature* 520. 23 April 2015: 431.

¹⁴ Sheperd, Jessica and Simon Rogers. "University guide 2012: download the Guardian tables and see how the rankings have changed." *Guardian* 17 May 2011. <http://www.theguardian.com/news/datablog/2011/may/17/university-guide-2012-data-guardian>

¹⁵ Daniel De Vise summarized this as follows: "Graduation and freshmen retention rates count for the largest share of the 100-point ranking, 27.5 percent, with extra points for schools whose graduation rates exceed statistical expectations based on socioeconomic mix and other factors. Academic reputation counts for 22.5 percent, based on surveys completed by presidents, provosts, admission deans and high school guidance counselors. Faculty resources (including student-faculty ratio) count for 20 percent. Selectivity (including admission rate) counts for 15 percent, financial resources for 10 percent, and alumni giving for 5 percent" (Vise, Daniel de. "U.S. News College Rankings Are Denounced but Not Ignored." *The Washington Post* 2 Sept. 2011.

This methodological self-reflexivity is of course better than its absence. But it should not lull us into a sense of rankings validity. Obvious issues include the following: there's quite a bit of reputational weight, there's a lot of emphasis on money, so richer schools do better, the data is mostly self-reported, *and* there's really nothing empirical in any of these surveys about student experience and learning, or about faculty and staff work life. There is no *qualitative* quality.

¹⁶ This image has been removed from the University of Tulsa website.

¹⁷ John Zilvinskis and Louis Rocconi, "Revisiting the Relationship between Institutional Rank and Student Engagement," American Educational Research Association <http://www.aera.net/Portals/38/Newsroom%20-%20Recent%20Research/Revisiting%20the%20Relationship%20between%20Institutional%20Rank%20and%20Student%20Engagement.pdf>

¹⁸ Nelson, Libby. "The US News Rankings Are Terrible for Students. Why Don't Colleges Stop Them?" *Vox*, September 9, 2014. <http://www.vox.com/2014/9/5/6106807/college-rankings-us-news-boston-clemson-problems>.

¹⁹ Symonds, Matt. "The Best Universities of 2013 - The Ranking of University Rankings." *Forbes*. N.p., 31 Dec. 2013. Symonds notes: "The Times Higher Education gives equal weight to teaching, research and citations, with a look at industry income and international outlook. Academic Ranking of World Universities puts a larger focus on research, accounting for 40% of the weighting. QS focuses on academic reputation 40%, employer reputation, student to faculty ratio, citations per faculty and international

factors. US News looks at the core missions of universities and groups them according to mission. They then look at employability, research, academic teaching quality, and they “test” the universities using academic polls and surveys.”

²⁰ Anon. “theunipod Ranking of University Rankings 2013 - US universities.”
<http://www.theunipod.com/making-your-choice/university-rankings/ranking-of-rankings/2013/us>

²¹ The introduction of non-standard factors do make a difference. *Washington Monthly* ranks schools by social factors such as the percentage of graduates without debt and the percentage that enter some kind of social service.

²² Shirky, Clay. “Your Massively Open Offline College Is Broken.” *The Awl*. 7 Feb. 2013.

²³ Desrochers, Donna M., Colleen M. Lenihan and Jane V. Wellman. *Trends in College Spending, 1998-2008*. Delta Cost Project, 2010. Figure 18.

<http://www.deltacostproject.org/sites/default/files/products/Trends-in-College-Spending-98-08.pdf>