"Faculty productivity" is hot. In part, that's due to what I think of as "the Texas wars." First came the skirmish at Texas Tech University in 2008, when faculty objected to a report commissioned by the chancellor on whether tuition was rising because faculty weren't in the classroom enough. Next was a controversial report last year on the Texas A&M system, comparing faculty salaries with the amount of money generated through teaching. Now the Regents' Task Force on University Excellence and Productivity has created its own metric, published recently online as an 821-page spreadsheet for the University of Texas system.

But the issue isn't limited to Texas. During the A&M uproar, a spokesman for the American Association of University Professors was quoted as saying that tough times are leading a number of states to look at faculty productivity. And it is important to remember that the topic is part of a larger public concern with accountability in higher education. This month the Miller Center of Public Affairs, at the University of Virginia, released a report calling on colleges "to focus on productivity." Echoing the message of the Obama administration, it noted that "at a time of budgetary stresses, colleges must be rewarded by both state and federal governments for producing more graduates."
isn't clear, however, how such productivity is to be measured. By the annual number of degrees awarded? Within how many years after matriculation?

The same kind of confusion is at the root of debates on faculty productivity.

Concern for faculty productivity actually goes back further than the recent focus on academic accountability. In the past, the questions were usually whether professors were overpaid, since they spend so few hours in face-to-face teaching. The comparison was to schoolteachers, who spend most of each day in the classroom. Of course, the faculty response has always been that this seemingly simple metric misses much of the real work of professors: out-of-class student contact, class preparation, research, and administrative duties.

By and large, the professoriate has carried the day, but its arguments have become less persuasive in recent years as the character of professorial work has been transformed. Research expectations have risen, along with pay, while teaching loads have fallen markedly, and well-paid outside consultations have proliferated.

An additional factor, insufficiently acknowledged, is the emergence of significant asymmetries in compensation according to the presumed market value of knowledge professed—some scientists and economists are extraordinarily better paid than scholars of language, literature, and the like (mainly in the humanities). Further, compensation (both salary and research support) in some fields derives from outside sources (government and industry). The result is that it is no longer clear that "professors" have a clear and unitary professional task that can be readily quantified.

Can we, then, measure our productivity? What the data mean so far—or how they will be used—is, to say the least, confusing. The report on Texas Tech concluded that faculty productivity was lagging, to which faculty responded that they weren't teaching less—the chancellor was trying to expand the university through competitive price-cutting. The A&M report found that faculty members were "generating more money than they are costing the university, although some of the most prestigious professors would appear to be operating in the red." The report noted that its data were for "management information only," but faculty were unsure what that meant.

In the report on the University of Texas system, each faculty member was listed with salary, the number of "sections" taught, and the research support brought in. We also learned the average grades they assigned students and their standard deviations, as well as average evaluation scores from students. But how do grades or student evaluations bear on productivity? The sparse materials accompanying the database did not explain. I haven't a clue.
Nor do the 821 pages shed much light on how the data will be used. There may, however, already be evidence that the report is being politicized: The task force is chaired by a member of the board of the Texas Public Policy Foundation, a conservative think tank closely allied with Gov. Rick Perry.

Just how fraught reducing education to numerical measures can be is reflected in yet another report, which has gone viral on the Internet this month. Richard Vedder, an economist at Ohio University and a *Chronicle* blogger, and his colleagues at the independent Center for College Affordability and Productivity found that 20 percent of faculty members on the flagship campus, at Austin, were teaching 57 percent of students' credit hours and generating 18 percent of external research funds; the least productive 20 percent were teaching 2 percent of credit hours and bringing in a smaller percentage of research funds than other faculty segments were. The lesson? Make the underperformers teach more, and gain productivity without losing research funds.

But as my colleagues the historians Anthony Grafton and James Grossman responded on a *New York Review of Books* blog, we know that large classrooms do not always produce the best educational experience, and that outside funds as a measure of anything vary by field.

Which takes us to the National Research Council’s rankings. Faculty have long obsessed over those assessments of doctoral programs. Most of us have long thought that those traditional reputational metrics have never found a reliable quantifiable way to measure quality; many of us now think that the most recent NRC measures are so obscure as to be nearly useless. Some very smart people have worked on the NRC rankings. If the NRC can’t do it, who can?

None of this is meant to dodge the issue. Because the faculty-productivity debate is part of the broader discussion of accountability in higher education, it is not going away. There are huge sums of public (and private) money invested in the higher-education industry. There is a near universal utilitarianism in the way our culture judges most everything, including our work. And it is not as though we faculty do not care about quality comparisons. We have long used very crude measures of productivity. (How many books and articles have you published? Where?) Most important, we have long complained about the impact of rising research expectations on our capacity to teach and advise students.

We have less frequently worried about the effectiveness of either our research or our teaching. We need to be able to explain what we do in a more thoughtful and precise manner, somehow formulating multiple measures of our diverse and complicated system of higher education. The first step is to describe our jobs clearly; then we ought to be able to determine how well we are doing them.
To do that, we must disaggregate the components and rethink how we can evaluate them separately. We are already developing ways to assess undergraduate learning outcomes, and we have reasonably reliable ways to assess scholarship (although not always or solely through quantification). Leadership programs and the like offer help when it comes to reviewing service and administrative duties. That will leave us the harder job of assigning importance to the various components, something we have traditionally avoided by collapsing everything into the assessment of research.

And, finally, we will have to admit that the assessments are relative to the types of educational institutions in which we work—research universities cannot and should not provide the implicit model for all of higher education.

Faculty have long resisted external assessments, and with good reason. Even paranoids have enemies. Still, we must be proactive in putting forth a better and more accurate way of assessing our complex and multiple work assignments.

*Stanley N. Katz directs the Center for Arts and Cultural Policy Studies at Princeton University’s Woodrow Wilson School.*