

Note: This essay was originally published as Smagorinsky, P. (2011, November 8). Efforts to "quantify teaching have become a national obsession on its way to becoming a national joke." *Atlanta Journal-Constitution*. Available at http://blogs.ajc.com/get-schooled-blog/2011/11/08/efforts-to-quantify-teaching-have-become-a-national-obsession-on-its-way-to-becoming-a-national-joke/?cxntfid=blogs_get_schooled_blog

Since then, the AJC has cleared out its archives, so it is now only available in the form in which I originally submitted it to Maureen Downey's Get Schooled education blog.

Teacherball by Peter Smagorinsky

The film *Moneyball* is based on Michael Lewis's book about Oakland Athletics' General Manager Billy Beane's application of Bill James's statistical tests for analyzing the productivity of baseball players. I read Bill James's annual *Baseball Abstract* back in the 1980s, and probably found it more interesting than I found actual baseball games. He was very insightful about challenging conventional wisdom about how to assemble players and manage them on the field. His statistical tests, for instance, found that a key measure of a hitter's performance, batting average, is deceptive because it masks other factors such as how often a hitter strikes out, hits into a double-play, or draws a walk to get on base. The point of an at-bat is thus poorly measured by one of the game's most cherished statistics.

The profession of teaching is currently subject to quantification for the purpose of identifying productive faculty members. I taught in high schools from 1976-1990, and much has changed since then. I taught in 3 high schools. Each relied on classroom visits by an administrator, who observed the instruction and then had a followup meeting to discuss the class with the teacher. Based on this observation, along with whatever other anecdotal evidence that came his or her way over the course of the year, the administrator provided an evaluation of the teacher that went into a personnel file. By and large, what mattered most was the quality of interactions with students and how the classes promoted learning, at least according to the expectations of the observing administrator.

Now, this system had its problems. At my first job, where I stayed for one year, the administrator took an immediate dislike to me and framed his evaluation as negatively as possible. At my second job, my department chair liked me and so saw the positive in my teaching. My third chair took a laissez-faire approach, letting faculty decide their own evaluation ranks. Because I always feel that I can do a better job, I always gave myself medium marks; while others on the faculty who were not respected by students always gave themselves the highest evaluations possible. The chair gave us what we asked for, without letting her judgment interfere. So, this fairly common system was not entirely reliable, if my experience of getting three very different levels of evaluation from three different administrators is any indication.

Today's world relies on different forms of evidence, particularly students' test scores, to evaluate, reward, and punish teachers. There is a key difference here between the world that Bill James changed and the world that Bill Gates and others are now influencing. James entered a field in which statistical data were used, but used poorly, and he introduced new analytic

measures that provided better, more nuanced, and more reliable information. The business model favored by today's educational policymakers introduces statistical measures to a field in which they have historically not been employed. Their primitive use of statistics, however, is more akin to the ill-informed world that Bill James sought to enlighten than the world that has now celebrated and adopted his view that complexity can be measured.

And the use of statistics to measure teaching performance is not simply primitive, but often ludicrous. I know of a school district in New Jersey where a teacher's performance is in part measured by the test scores of students who take standardized tests that the teacher is proctoring. In all likelihood, the teacher hasn't taught these students; rather, her role is to walk around the room making sure that nobody's cheating. [Valerie Strauss also describes the "nutty" approach of rewarding teachers for running into parents while shopping as evidence of their engagement with parents.](#) I'm guessing that we could find quite a few additional misguided evaluation approaches with enough ferreting, and such a project would be worthwhile.

OK, these measures are so dumb that anyone beside the people implementing them would find them laughable. Standardized test scores strike me as being similar to measuring batters by their batting average while overlooking other qualities of a turn at bat in the context of an inning or game. Sometimes you need one run, sometimes you need five; sometimes there's a runner on base and sometimes there's not; and what you do at the plate is a function of the context of the at-bat. A strikeout and a ground ball that advances a runner look the same on a batting average, but their impact on the game is very different.

The problems surrounding standardized tests have been so well hashed over that I hope I needn't rehash them here. They are not meaningless; but they hardly tell the whole story, or even a good part of it. I am not a statistician, and so am in no position to assume the role of Bill James for Education. Apparently, however, no one else is either, and thus the efforts to quantify teaching have become a national obsession on its way to becoming a national joke.

And yet, few people, including teachers, believe that teaching cannot and should not be evaluated in any way. Most teachers who stay in the profession are dedicated enough to find their incompetent colleagues to be an embarrassment. There needs to be a way to distinguish between the two in such a way that the effort does not have the "guilty till proven innocent by means of questionable measures" quality of the current quantification mania. I have contended on a number of occasions that the best way to assemble and retain a quality faculty is to provide quality work conditions. A punitive environment designed to intimidate faculty will not produce a workplace where great teachers want to be.

Are statistics the answer? If so, they need to provide something more complex than the crude measures currently being employed: test scores that teachers rightfully find misplaced and disrespectful toward the difficult work that they do. Bill James moved baseball forward by getting behind the superficiality of conventional wisdom and finding out which factors predict the scoring and prevention of runs, which in turn predict winning. Often this effort involved looking where nobody else had looked before; often it involved adding new variables to existing means of measurement. Each was designed to identify with some precision what produced success.

Teaching and learning are a lot more complicated than baseball. In baseball, there's one outcome that everyone agrees on: winning games by outscoring opponents. Educational goals are much more diffuse. In addition to being responsible for students' learning of subject matter, however determined and measured, teachers are charged with providing substance abuse education, giving attention to learning variation based on both biological and cultural factors, ensuring school safety, supervising extracurricular activities, designing the curriculum, developing compassionate and engaged citizens, preparing students for both college and the workforce, providing paper trails for each students' progress or lack thereof, inputting data that comprise students' records, acting *in loco parentis* for the many students whose parents are not present in their lives, policing student behavior, providing counseling to wayward or confused students, teaching the application of abstract concepts, providing character education, working effectively with parents, working productively with colleagues and administrators, and doing much more; and further, they are responsible for preparing students for the standardized tests that have nothing to do with any of the above, yet provide the single score by which their teaching is evaluated.

Surely we can do better than that, even as Arne Duncan insists that it's the only way to go. There's an old saying in statistics that if you can't measure something, measure it anyway. That's what we're doing now, and frankly it's a disaster for teachers and students. Is this a profession that you want to enter and remain in under these conditions? If not, how can you help to contribute to a system that good teachers will buy into because it helps to make schools a more emotionally satisfying and intellectually challenging place them and their to spend a good portion of their lives in?