Journal of Applied Research on Children: Informing Policy for Children at Risk

Volume 3
Issue 2 *Measuring Success in Public Education*

Article 20

2012

The Problem with Value-Added Measurement

Gary Rubinstein
Stuyvesant High School, garyrubinstein@yahoo.com

Follow this and additional works at: http://digitalcommons.library.tmc.edu/childrenatrisk

Recommended Citation
Rubinstein, Gary (2012) "The Problem with Value-Added Measurement," *Journal of Applied Research on Children: Informing Policy for Children at Risk*: Vol. 3: Iss. 2, Article 20.
Available at: http://digitalcommons.library.tmc.edu/childrenatrisk/vol3/iss2/20

The Journal of Applied Research on Children is brought to you for free and open access by CHILDREN AT RISK at DigitalCommons@The Texas Medical Center. It has a "cc by-nc-nd" Creative Commons license" (Attribution Non-Commercial No Derivatives) For more information, please contact digitalcommons@exch.library.tmc.edu

The modern "accountability movement" of school reform requires schools and teachers to be constantly rated and compared. Though schools and teachers have always been evaluated—teachers by their supervisors, and schools by higher-level administrators—reformers have declared these evaluations broken. Their rationale is that many students are not passing their standardized tests. The teachers and the schools responsible for this clear case of educational neglect, they say, have never been punished appropriately. Schools, up until recently, have not been closed down for poor performance. Under-performing teachers have been spared the pink slips they so deserve.

The reason that "failing" schools and "ineffective" teachers have managed to get satisfactory ratings, they continue, is that evaluations have been based solely on measuring somewhat subjective inputs, like what the teacher and school can be observed doing, rather than objective outputs, which are the result of those teacher and school inputs—the amount of student learning that occurs.

This suggests that there is some percentage of teachers in this country that, despite great effort, is not really accomplishing anything. What percent this is, is never defined, but some reformers quote the work of Stanford's Eric Hanushek, who says that schools are wise to follow the philosophy of G.E.'s Jack Welch and annually fire the bottom 5% of workers each year, as defined by these outcomes.[1,2] In certain cities, mayors seem to believe that annually closing the bottom 5% of schools also will lead to higher outcomes.[3]

There are certainly professions where the outcomes are easy to measure. A fisherman, for example, either catches a lot of fish or he doesn't. If he catches no fish despite having the appropriate inputs (he gets to the spot early, has the proper bait, spends all day on the boat), he is not going to get any money when it is time to sell the evidence of his daily work. Perhaps he came very close to catching many fish, but nobody wants to buy a can of "almost-tunafish." He's not getting it done and needs to find a new line of work.

The dream is that we can accurately measure the outcomes of teaching in a similarly efficient and fair way. Economists have spent twenty years chasing this elusive measuring stick. And though they haven't even come close to developing it, the self-anointed education reformers don't have time to wait. So they have coerced states across the country, as a requirement to be eligible for Race To The Top money, to measure teacher and school quality through something called "value-added" measurements.

The idea behind "value-added" measurements is that a computer

can take all the information about a class of students and determine what those students would score on the state exams ten months from now if they were taught in an "average" school by an "average" teacher. Ten months later the students take the test, and based on whether their results exceed or fall short of the computer's prediction, that school is given an "A" or an "F" rating and the teacher is correspondingly rated as effective or ineffective.

Unfortunately for schools and teachers, the computer isn't very good at predicting what it is supposed to. One way to check if a measuring tool is useful is to see if it is consistent. So if you have a scale and one day it says you weigh 140 pounds and the next day it says you weigh 300 pounds, well, this isn't a scale you want your livelihood to depend upon. Yet this is exactly what happens with value-added. A teacher rated highly effective one year might be highly ineffective the next, despite, according to the teachers, the fact that they didn't do anything wildly different.[4] When teachers and schools are forced to play Russian Roulette this way, it is not good for teachers, it is not good for schools, and, no, it is not good for students.

One issue is that the state tests are not very good. Reformers say that they are working on this issue. And while I'd also like to see better tests, I don't think that it is the best use of our money. If we spend a trillion dollars developing the ultimate tests that finally do identify all the "impostor" teachers who would have slipped through the input-based evaluation system otherwise, all we will learn is that there are not so many of them. We will have wasted all that money to confirm what should be common sense to anyone who has ever been a teacher, a parent, or a student. Threatening a teacher with termination or a school with turnaround unless they measure up on an unreliable scale is never going to improve education.

I'm not convinced that a faulty measure of outcomes is better than an accurate measure of inputs. Neither measures student achievement, student learning, or student growth directly, but the inputs are more closely correlated, given the crude state of the value-added metrics today. Until outcomes can be measured more accurately, we need to stick with measuring the inputs we can.

Even the most extreme reformers admit that the way we currently measure outcomes is a work-in-progress, which is why they are sure to tell you that they think evaluation should be based on multiple measures. But then the issue is, what percent should the value-added measure be of the entire rating? When the D.C. IMPACT model was created a few years ago, the value-added component was set at 50% for teachers in tested

subjects.[5] After a few years and several hundred fired teachers, researchers started looking at the data and noticed that the value-added component did not agree with, in general, the principal evaluation component.[6] As a result of this, they have now reduced the value-added to 35%.[7] Years from now it will probably drop even more.

One of Secretary of Education Arne Duncan's pet phrases when it comes to defending the current state of test-based accountability is, "We always let the perfect be the enemy of the good in education," suggesting that this current way of attempting to measure outcomes is good. But I'd like to answer back to him that, "'The perfect' CAN be the enemy of 'the crummy'." Changing an imperfect system for one that is even more imperfect is not reform.

References

1. Hanushek EA. *Teacher Deselection.* http://www.stanfordalumni.org/leadingmatters/san_francisco/documents/Teacher_Deselection-Hanushek.pdf. Accessed October 3, 2012.
2. DiCarlo M. Will firing 5-10 percent of teachers make us Finland? *Washington Post.* December 18, 2010. http://voices.washingtonpost.com/answer-sheet/research/will-firing-5-10-percent-of-te.html. Accessed October 3, 2012.
3. New york city school turnaround plan to close schools, fire teachers blocked by judge. *Huff Post Education.* July 25, 2012. http://www.huffingtonpost.com/2012/07/25/new-york-city-school-turn_n_1703302.html. Accessed September 20, 2012.
4. Clifford T. Charting the stages of teacher data report grief. SchoolBook.org Website. http://www.schoolbook.org/2012/03/09/charting-the-stages-of-teacher-data-report-grief/. Accessed October 3, 2012.
5. Sawchuk S. D.C. unveils complex evaluation system. *Education Week.* October 5, 2009. http://blogs.edweek.org/edweek/teacherbeat/2009/10/dc_unveils_complex_evaluation.html. Accessed October 3, 2012.
6. Turque B. Link between value-added ratings and classroom observations is modest at best, studies show. *Washington Post.* April 27, 2012. http://www.washingtonpost.com/blogs/dc-schools-insider/post/link-between-value-added-ratings-and-classroom-observations-is-modest-at-best-studies-show/2012/04/27/gIQANkJulT_blog.html. Accessed October 3, 2012.
7. District of Columbia Public Schools. Chancellor Henderson announces updates to teacher evaluation system. http://dcps.dc.gov/DCPS/About+DCPS/Press+Releases+and+Announcements/Press+Releases/Chancellor+Henderson+Announces+Updates+to+Teacher+Evaluation+System. Published August 3, 2012. Accessed October 3, 2012.