



MEASURING TEACHER EFFECTIVENESS: Credentials Unrelated to Student Achievement

Marcus Winters, *Senior Fellow*

Given the challenges facing American public education today, identifying effective teachers is a more vital task than ever before. A wide body of research shows that teachers are the most important school-based factor related to student achievement. Policymakers and taxpayers want to know what factors create effective teachers—not only for the sake of their own children’s educations but also because teacher salary and benefits represent the nation’s single largest educational expenditure. And school administrators need to identify teachers who will be successful over the long term before those teachers earn the ironclad job protection of tenure.

In the U.S. public school system today, the method used to determine teacher effectiveness—and thus to drive salary, promotion, and tenure decisions—is based on a few external credentials: certification, advanced degrees, and years of experience in the classroom. Yet according to a new analysis of student performance in Florida that two colleagues and I conducted, little to no relationship exists between these credentials and the gains that a teacher’s students make on standardized math and reading exams. Our expansive study included all test-taking public elementary school students in the state of Florida over a period of four years.

Our study, to be published in the peer-reviewed journal *Economics of Education Review*, builds on two decades of research from a

variety of school systems and confirms a consistent finding: external teacher credentials tell us next to nothing about how well a teacher will perform in the classroom. Such research has not, however, yet had a substantial effect on the practices of U.S. public schools. Today's public school system continues to rely on external teacher credentials to decide who gets to teach and how much a teacher is paid. Though the debate over how most accurately to use statistical measures to identify teacher quality is far from completed, the general finding that there is a vast difference between the system's best and worst teachers is no longer in serious dispute. The large body of research on teacher quality suggests that a new method of identifying the best teachers is needed—one that focuses on measuring the contributions that teachers actually make in the classroom.

THE CURRENT SYSTEM

In making decisions on pay, promotion, and tenure, U.S. public schools today do not seriously consider measures of how well a teacher performs in the classroom. Instead of distinguishing between the observed performances of teachers, the current system differentiates teachers by the number of advanced degrees that they hold and their years of experience in the classroom.

Teachers must acquire a certification, the vast majority of which are gained by graduating from a college of education. Once a candidate earns a teaching position, however, that teacher no longer undergoes serious evaluation. The evaluation system relies on infrequent and superficial observations of a teacher's performance, often during a single class period (or less) at a time known by the teacher in advance. Consequently, nearly all teachers are classified as effective: an analysis of 12 large school districts in four states by the New Teacher Project found that in districts using a binary evaluation system (i.e., with only two ratings: "satisfactory" or "not satisfactory"), over 99 percent of teachers received the thumbs-up rating. Even districts that used broader evaluation distinctions ranked 94 percent of teachers in one of

the top two tiers of effectiveness and deemed just 1 percent "unsatisfactory."¹

The vast majority of public school systems pay teachers uniformly, according to a salary ladder built on only two key factors: advanced degrees and experience. In American public schools, the single-salary schedule has a long history that actually predates the rise of teachers' unions as a political force: as early as 1918, about 65 percent of American urban school systems reported that they had established a salary schedule.² Today, single-salary schedules based entirely on these two factors are the norm nationwide.

THE FLORIDA STUDY

Along with my coauthors Jay Greene and Bruce Dixon (professors at the University of Arkansas), I utilized a data set provided by the Florida Department of Education to study the relationship between elementary school teachers' training and experience, on the one hand, and the learning gains made by their students in a given year, on the other.

We used the data to study the relationship between very detailed information about teachers' credentials and the achievement of their students. Essentially, our data matched public school elementary students in Florida to their teachers over several years and measured whether a student's performance improved when the student was assigned to a teacher with a particular set of credentials. For a subsample of teachers who were educated in a Florida public university, we were able to measure the relationship between the types of courses that the teachers took in college and the achievement of their students. The Florida data set is particularly rich in that it follows all test-taking students over a sustained period of time, matches students to individual teachers, and includes information on the number and type of college courses taken by teachers who graduated from a Florida public university.

As with most previous research, we found *no* relationship between a teacher's earning a master's degree,

certification, or years of experience and the teacher's classroom performance as measured by student test scores. Though we found that some pedagogy course work was related to teaching effectiveness, the magnitude of the effect was mild: even very detailed information about the teacher's preparation in college told us very little about how effective that teacher would be in the classroom.

CONFIRMING PREVIOUS RESEARCH

Our results from Florida only confirm a wide body of existing research. Not a single one of the 34 studies that used a "high-quality" methodology (i.e., methodology that accounted for previous student test scores) evaluated in a recent review of the research by Eric Hanushek and Steven Rivkin found a relationship between a teacher's earning a master's degree and student achievement.³

Empirical research on the effect of classroom experience yields more complex results than research on teacher credentials; but ultimately, it is just as discouraging. Most research does find a positive relationship between the number of years a teacher has spent in the classroom and his or her influence on student achievement—but *the benefit of that experi-*

ence appears to plateau after the third to fifth year. The average teacher is at his or her worst during the first year in the classroom, gets better in the second year, a little better in the third year, and then never gets any better after that.⁴

Research suggests that 3 to 5 percent of teachers' contribution to their students' math and reading scores can be explained by factors that are typically collected in an administrative data set⁵—that is, professional credentials and years of experience. Upward of 97 percent of what makes one teacher more effective than another is unrelated to factors such as the number of years the teacher has been teaching and the credentials that the teacher has earned.

CONCLUSION/RECOMMENDATIONS

Modern research on teacher quality makes clear that the factors used to determine a teacher's compensation tell us little to nothing about how well the teacher will perform in the classroom. That consistent finding has (or should have) enormous implications for the future of the current system. The results of an employment policy based entirely on credentials uncorrelated to student achievement are obvious: we see wide variation in the quality of public school teachers.

NOTES ON METHODOLOGY

Our study attempted to improve upon previous research in this area by accounting for the fact that a large number of teachers leave the profession within their first few years of teaching. If such attrition is related to the teacher's effectiveness in the classroom—that is, perhaps the worst teachers are most likely to exit the classroom—then previous estimates of the relationship between teacher characteristics and student achievement could be biased. A primary motivation for our study was to evaluate whether accounting for such "sample selection" directly alters the estimated relationship between external teacher credentials and student achievement.

Though our results show that sample selection is present when estimating the relationship between observed teacher characteristics and student achievement, accounting for it statistically does not substantially influence the estimates. The important policy implication of that finding is that it confirms that sample selection did not seriously bias previous estimates of the relationship between teacher characteristics and student achievement.

The structure of the current system is simply indefensible, given modern research findings. There is nothing inherently wrong with relying on proxies for effectiveness when making employment decisions. However, when those proxies fail to differentiate meaningfully between the most and the least productive workers, they should be jettisoned. This is certainly the case with our public schools, where wide variation in teacher quality persists among those who have passed through the usual screens and earned the recommended degrees.

Just how much does teacher quality vary? An early study by Stanford University economist Eric Hanushek estimated that the difference between being assigned to one of the system's best teachers and one of its worst is about an additional grade level's worth of proficiency at the end of the school year.⁶ Similar variations in teacher quality have been found in Tennessee,⁷ New Jersey,⁸ Chicago,⁹ Florida,¹⁰ and other unnamed school districts across the country.

Most parents, teachers, and school administrators surely won't find it surprising that teachers are not identically effective. What to do?

The findings of our research and other studies suggest that public schools should revise their employment systems so that they prioritize measures of the teacher's actual effectiveness in the classroom. School systems should develop comprehensive evaluation systems that utilize quantitative (e.g., test scores) as well as qualitative (e.g., classroom observation) measures of teacher effectiveness. The results of these evaluations should be used to determine which teachers are retained in the classroom, how much a teacher is paid, and whether the teacher receives job protections in the form of tenure.

Though there is much work to be done, several school systems have begun developing and adopting rigorous evaluation systems for their public school teachers with the hope of using their results to inform employment decisions. Over the last several years, researchers have used variants of a technique known as "value-added" analysis to study the extent to which teacher quality varies. Essentially, this procedure measures how much an individual teacher contributes to his or her student's test scores after accounting for factors that the child brings into the classroom. Research utilizing the value-added approach consistently finds that teacher quality is the most important factor within a school's control for boosting student performance. Students with nearly identical backgrounds will perform quite differently on standardized tests, depending on the teacher to whom they were assigned.

Taking into account measures of a teacher's effectiveness has an enormous effect on the public school system. For instance, New York City's use of a new evaluation system when making tenure decisions this year led to a dramatic decrease in the number of teachers receiving those job protections. School systems across the country should continue moving toward emphasizing direct measures of teachers' performance, rather than credentials, to identify and compensate the best teachers.

Over the last two decades, we have learned two important lessons about public school teachers: teacher quality varies dramatically; and almost nothing we know about a teacher before he or she enters the classroom accurately predicts how successful that teacher will be. Now heavily documented through empirical research, these findings should point us toward a fundamental transformation of our system for evaluating public school teachers.

ENDNOTES

- ¹ D. Weisberg et al., “The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness” (New Teacher Project, 2009).
- ² E. S. Evenden, “Teachers’ Salaries and Salary Schedules in the United States, 1918–19,” paper prepared for the Commission on the Emergency in Education of the National Education Association (1919), 119.
- ³ E. A. Hanushek and S. G. Rivkin, “Teacher Quality,” in Eric Hanushek and Finis Welch eds., *Handbook of the Economics of Education*, vol. 2 (Amsterdam: North-Holland, 2006), pp. 1051-1075.
- ⁴ For a review of this research, see *ibid.*, pp. 1058-1062.
- ⁵ D. Goldhaber, “The Mystery of Good Teaching,” *Education Next* (spring 2002): pp. 50-55.
- ⁶ E. Hanushek [need title, etc.], 1992.
- ⁷ W. L. Sander and J. C. Rivers, “Cumulative and Residual Effects of Teachers on Future Student Academic Achievement,” University of Tennessee Value-Added Research and Assessment Center (1996).
- ⁸ J. E. Rockoff, “The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data,” *American Economic Review* 94, no. 2 (2004): 247–52.
- ⁹ D. Aaronson, L. Barrow, and W. Sander, “Teachers and Student Achievement in the Chicago Public High schools,” *Journal of Labor Economics* 25, no. 1 (2007): pp. 95-135.
- ¹⁰ M. West and M. Chingos, “Teacher Effectiveness, Mobility, and Attrition in Florida: A Descriptive Analysis,” paper prepared for *Performance Incentives: Their Growing Impact on American K–12 Education conference* (2008).