Examining the Relationship between Student Achievement and Observable Teacher Characteristics: Implications for School Leaders*

Anna Jacob

Abstract

The literature on effective schools emphasizes teacher quality as a critical input for improving student outcomes. Teacher certification status, years of experience and possession of a graduate degree are three factors commonly taken into consideration when assessing teacher quality. With the advent of advanced measurement tools, however, principals and other school leaders have an additional means to assess teacher quality. Rigorous evidence is assessed to determine the potential merits of value-added measurement tools to assist principals in developing teacher talent and making staffing assignment, development, and retention decisions.

1 NCPEA Publications

NOTE: This manuscript has been peer-reviewed, accepted, and is endorsed by the National Council of Professors of Educational Administration (NCPEA) as a significant contribution to the scholarship and practice of education administration. In addition to publication in the International Journal of Educational Leadership Preparation,[1] Volume 7, Number 3 (Winter 2012), ISSN 2155-9635, this manuscript exists in the Connexions Content Commons as an Open Education Resource

*Version 1.4: Oct 11, 2012 8:55 am -0500
†http://creativecommons.org/licenses/by/3.0/
[1]http://www.ncpeapublications.org/latest-issue-ijelp.html

(OER). Formatted and edited by Theodore Creighton, Virginia Tech; Brad Bizzell, Radford University; and Janet Tareilo, Stephen F. Austin State University. The assignment of topic editor and double-blind reviews are managed by Editor, Linda Lemasters, George Washington University. The IJELP is indexed in the Education Resources Information Center (ERIC), sponsored by the United States Department of Education under Contract No. ED-04-CO-0005.

2 Sumario en español

La literatura en escuelas efectivas le acentúa a maestro calidad como una entrada crítica para mejorar resultados de estudiante. El estatus de la certificación del maestro, los años de experiencia y posesión de un grado graduado son tres factores tomaron en consideración comúnmente al valorar calidad de maestro. Con la venida de herramientas avanzadas de medida sin embargo directores y otros líderes de escuela tienen un medios adicionales para valorar calidad de maestro. La evidencia rigurosa es valorada para determinar los méritos potenciales de herramientas de medida de valor agregado a ayudar a directores en el talento revelador de maestro y haciendo tarea que provee, el desarrollo, y las decisiones de retención.

NOTE: Esta es una traducción por computadora de la página web original. Se suministra como información general y no debe considerarse completa ni exacta.

3 Introduction

"And this is our present purpose: to discover, so far as possible, what elements enter into the making of a capable teacher." - J.L. Meriam, Teachers College Contributions to Education No. 1 (1906)

School administrators are well aware that teachers matter for student achievement. It is also generally recognized that there is wide variation in teacher effectiveness both within and between schools (Kane, Rockoff, & Staiger 2008; Lankford, Loeb, & Wyckoff 2002; Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004). As recently as two decades ago, we had no way of quantifying the variation in teacher classroom effectiveness. Many concluded that while teachers are hugely important, variation in teacher effectiveness within and between schools is based on unobservable teacher characteristics that are difficult if not impossible to measure. School administrators could recognize the difference in learning outcomes that would occur when a child was assigned to one teacher over another but quantitatively measuring the extent of a teacher's effectiveness was a daunting challenge. Luckily, the empirical revolution that has occurred in education over the past few years has led to the development of a set of tools that are capable of quantifying the extent of variation in teacher effectiveness for the first time (McCaffrey, Lockwood, Koretz, & Hamilton, 2003). Although imperfect, tools such as "value-added" measuring provide us with useful estimates that help describe the variation in teacher effectiveness. "Value-added" methods are a collection of complex statistical techniques that isolate teacher and school contributions to student learning from their students' backgrounds and preparation. These new statistical methods improve upon traditional metrics that describe average student achievement at a school by capturing student achievement *gains*, conditional on student background and preparation, instead of measuring student achievement *levels*, which could be strongly influenced by student family disadvantage or other background characteristics.

When judging the value of any new tool it is essential to consider the counterfactual- if we dislike "value-added" measurements, what are the alternative approaches to measuring teacher quality? This paper examines three common measures that school leaders may take into consideration when assessing teacher quality and compares the usefulness of these measures with more recently developed "value-added" methodologies.

The next section establishes the need for accurate measures of teacher quality by describing evidence on the importance of teachers for both short and long term student outcomes. The following three sections contain a comprehensive review of the prodigious body of research related to the relationship between teacher experience, graduate degrees, and licensure test results on one hand and student achievement on the other. The specific research question addressed is whether these observable characteristics of teachers are associated

with high student outcomes. The final section concludes by describing the potential of value-added measuring tools to assist school leaders when making decisions about teachers.

3.1 Teachers Matter

School leaders have long been aware of the potential impact of an excellent teacher on student achievement outcomes. Recent developments in analytic methods have allowed us to begin to quantify that impact for the first time (Buddin, 2011; McCaffrey et al., 2003; Sanders & Horn, 1998). Large-scale, longitudinal studies in Texas and New Jersey provide robust evidence of the extent to which teachers matter for student academic success.

Researchers Rivkin et al. (2005) use a panel data set of student outcomes on the Texas state test spanning grades 3 to 7 for three cohorts of students in the mid-1990s to identify sources of differences in student achievement, providing strong evidence that a one standard deviation increase in teacher quality (measured in terms of prior student achievement gains on the state test) can have a 0.10 standard deviation impact on student achievement in math and a slightly smaller effect in reading. Rockoff (2004) affirms this finding by using a random-effects meta-analytic approach with New Jersey data for two contiguous districts from 1989-90 through 2000-01 to measure the variance of teacher fixed effects on student achievement. Rockoff's empirical results indicate large differences in teacher quality within schools, concluding that a one standard-deviation in teacher quality raises test scores by approximately 0.10 standard deviations in reading and math.

The magnitude of teacher impacts on student achievement outcomes is once again affirmed by findings using data for Chicago public high schools by Aaronson, Barrow, and Sander (2007). The measure of teacher quality employed in this study is the effect on ninth-grade math scores of a semester of instruction with a particular teacher, controlling for prior-year math scores and a range of observable student characteristics. The authors find a one-semester improvement in math teacher quality raises student math scores over one year by approximately one-fifth of average yearly gains. The magnitude of this estimate is statistically similar to the results reported by Rockoff (2004) and Rivkin et al. (2005).

Hanushek (1992) has quantified the difference between having the "best" and the "worst" teacher for one school year using data that were generated over a 4-year period by the Gary Income Maintenance Experiment, incorporating test scores from the Iowa Reading Comprehension and Vocabulary tests. He ranks all teachers based on effectiveness measured by student test scores and calculates the difference between being assigned to a teacher at the twenty-fifth percentile as compared to being assigned to a teacher at the seventy-fifth percentile in quality, finding that the difference is about an additional grade-level's worth of proficiency by the end of the school year. Studies by researchers in Tennessee (Sanders & Rivers, 1996), New Jersey (Rockoff, 2004), Chicago (Aaronson et al., 2007) and Florida (West & Chingos, 2008) have all reached similar conclusions.

As implied by the consensus reached by these independent studies of the magnitude of teacher impacts, most contemporary education researchers agree that teachers matter (Angrist & Lavy, 2001; Goldhaber & Brewer, 1997; Jacob & Lefgren, 2004; Jepsen & Rivkin, 2002; Kane, Rockoff, & Stagier, 2006; Kane & Staiger, 2005; Rivers & Sanders, 2002; Rockoff 2004). The biggest issue facing school leaders, researchers and others in the education community is quantifying how much particular teacher characteristics matter and whether it is possible to predict teacher performance based on characteristics observed at the time of hiring. The next section attempts to synthesize the findings on three tangibles relating to measurement of teacher quality.

4 Measuring Teacher Quality

Traditionally, school leaders have had a limited number of indicators with which to judge teacher quality. This can be particularly problematic in schools serving low-achieving, poor, and minority students, which often face difficulties attracting and hiring effective teachers (Ferguson, 1998; Krei, 1998; Lankford et al., 2002). Because schools such as these do not attract as many applications from high quality candidates,

principals would greatly benefit from tools that help them accurately identify the highest quality applicants. Three common indicators of teacher quality that principals and school leaders currently rely upon are teacher experience, possession of graduate degrees, and teacher certification. In what follows, I present the research results judging the efficacy of these three indicators and conclude with a recommendation for a fourth indicator.

4.1 Teacher Experience

How much does teacher experience matter? It is intuitively appealing that a teacher's worst year is most likely his/her first year in the classroom, when classroom management issues are being tackled for the first time, organizational routines are being established, curricula mapped, procedures developed, etc. As a teacher's experience grows, however, we might expect that he/she is more likely to have figured out the most appropriate responses to a variety of classroom situations and problems, developed a strong sense of self-efficacy, and that student achievement increases as a result (Tschannen-Moran & Woolfolk Hoy, 2001; Tschannen-Moran, Woolfolk Hoy, & Hoy, 1998; Woolfolk & Hoy, 1990; Woolfolk Hoy, Hoy, & Davis, 2009). Indeed, the way that current teacher compensation systems are set up reveals a belief that experience correlates with effectiveness. Teachers are rewarded with salary increases for additional years of service and advanced degree attainment, not for other factors such as student achievement, teaching in hard-to-staff schools, or teaching a hard-to-staff subject such as math or science. Although the majority of current systems are set up to reward teacher experience, what does the research literature actually tell us about the predictive power of teacher experience on student achievement outcomes?

One of the inherent difficulties associated with measuring the correlation between teacher experience and student achievement outcomes is the bias that statistical models are prone to in the face of selection effects. There is evidence that better trained and more experienced teachers are assigned students of greater ability and students who display fewer discipline problems (Clotfelter, Ladd, & Vigdor, 2006). Given this selection effect, most regression models would not be able to control for all of the student characteristics that result in positive matching between student quality and teacher experience, resulting in upwardly biased estimates of the effects of teacher experience on student outcomes.

Researchers have three options in the face of selection bias. The first and strongest design is to implement a randomized, controlled trial which randomly assigns students to teachers with various levels of experience, ensuring that any differences in student characteristics, observed or unobserved, is due to chance and thus cannot systematically bias findings. This experimental design is considered the gold standard because "randomized field trials provide the most reliable answers" (Whitehurst, 2012). RCTs are expensive and often impractical to implement, however, which is why only a limited number exist in the field of education despite a commitment by IES to use randomized field trials whenever possible (Whitehurst, 2012). The second option, therefore, is to take advantage of a "natural experiment"- that is an exogenous change in student assignments to teachers of different experience levels that occurs naturally and is not correlated with student achievement. Unfortunately, natural experiments are not very common, which is why most research in this area has relied upon the third option. Aided by the development and availability of longitudinal student-level databases in recent years, researchers have employed student fixed effects models which can account for unobserved student characteristics, thus removing the bias associated with selection effects related to student sorting. Research by Rivkin et al. (2005) in Texas, for example, employs student fixed effects in the model specification to measure teacher impacts, reporting that "the learning curve appears to be quite steep in the first year or two of teaching before flattening out" (p. 435). Similarly, Rockoff (2004) uses student fixed effects in his models so that variation in student characteristics that are unchanging over time does not influence estimates of the differences in student performance across teachers.

I have described three strategies for handling selection bias relating to students but Rockoff (2004) and Kane, et al. (2006) draw our attention to the related problem of sample bias as it relates to teacher attrition from a study sample. If, for example, less effective teachers are more likely to leave the profession early, a study of the impact of teacher experience might falsely conclude that experience raises student outcomes when in reality, student achievement is rising because less effective teachers are exiting the profession. The

bias could work in the opposite direction too if highly effective teachers leave the profession to pursue more challenging, higher compensated opportunities, leaving researchers to spuriously identify a negative relationship between teacher experience and student achievement outcomes. Harris and Sass (2008) describe a technique to deal with this issue. By including a teacher fixed effect, researchers can control for unobserved teacher ability, resulting in unbiased estimates of the impact of teacher experience. Studies that address these methodological concerns in their empirical approach provide the most reliable evidence of the impact of teacher experience on student achievement outcomes.

Hanushek, Kain, O'Brien, and Rivkin (2005), Kane et al. (2006), and Rockoff (2004) take this approach in their models, in most cases finding modest effects of experience limited to the first few years of a teacher's career, suggesting that teacher effectiveness grows in the initial four or five years in the classroom and then begins to level off. Results for Vocabulary, Math Computation, and Math Concepts in Figures 1 and 2 demonstrate the growth in teacher effectiveness from approximately years one through five, followed by a plateau effect from around year six onwards. In just one case, Rockoff (2004) found that the marginal returns to experience exhibit much slower declines for Reading Comprehension but cannot rule out a violation of his identification assumption (Figure 1).

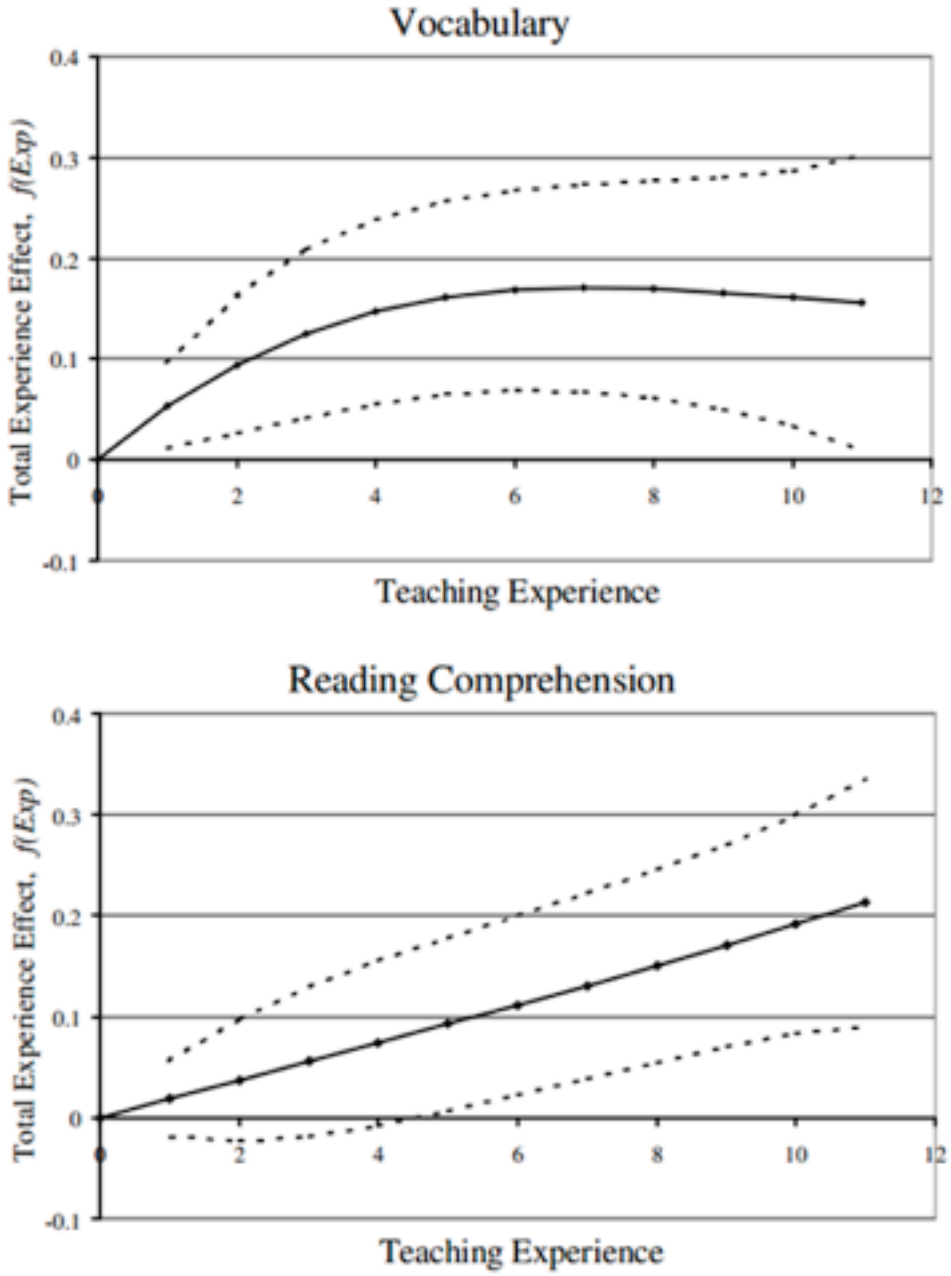

Figure 1. The Effect of Teacher Experience on Reading Achievement, Controlling for Fixed Teacher Quality. Dotted lines are bounds of the 95% confidence interval. Data shown is for two contiguous New Jersey school districts covering a period of up to twelve years Rockoff (2004).

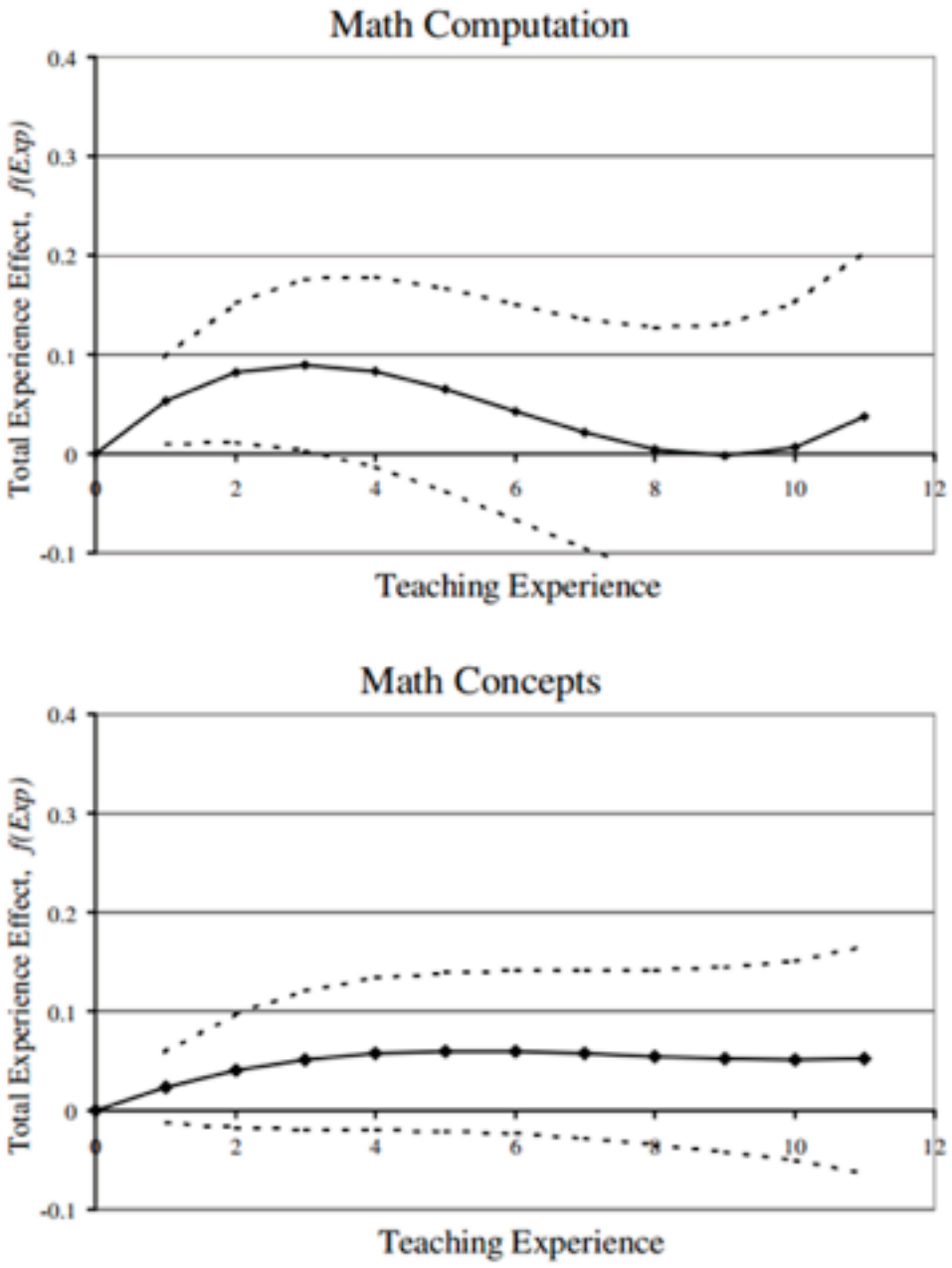

Figure 2. The Effect of Teacher Experience on Math Achievement, Controlling for Fixed Teacher Quality. Dotted lines are bounds of the 95% confidence interval. Data shown is for two contiguous New Jersey school districts covering a period of up to twelve years (Rockoff, 2004).

Research by Staiger and Rockoff in the nation's two largest school districts, New York and Los Angeles also address the methodological concerns outlined above when they incorporate teacher fixed effects in their models to measure the difference in average teacher effects on student achievement during their first few years' teaching as compared to the effects of experienced teachers. In both cities, results are consistent with Hanushek et al. (2005), Kane et al. (2006), and Rockoff (2004): student achievement rises rapidly during a teacher's initial years on the job then quickly flattens out. Similar findings have been reported in Texas

(Rivkin et al., 2005), North Carolina (Clotfelter et al., 2006), and Florida (Harris & Sass, 2006). For a general review of the literature, see Jacob (2007). Assignment to a teacher in her first year on the job results in average student achievement gains .06 to .08 standard deviations of achievement below likely achievement with a more experienced teacher. The magnitude of this difference shrinks in a teacher's second year on the job, however, with average student achievement gains .01 to .04 standard deviations below the gains for similar students assigned to more experienced teachers. In Kane and Staiger's 2008 study of the Los Angeles Unified school district, assignment to a teacher in her third year resulted in equal gains as assignment to a more experienced teacher while a small difference persisted in student achievement gains for third-year teachers in New York (Kane et al., 2008). To summarize, teachers' effectiveness rises rapidly in the first few years on the job but any benefits accrued by teacher experience plateau quickly. Administrators seeking a measure of teacher quality should incorporate more reliable indicators that can more accurately differentiate between teachers.

4.2 Graduate Degrees

Schools and district leaders seeking to raise student performance often seek to attract teachers with advanced degrees and to encourage existing teachers to pursue higher degrees in the hope that the process of acquiring such qualifications leads to a deeper understanding of teaching, learning, and human development, which in turn should lead to improved student performance. Ideally, a master's degree in education or similar qualification signals advanced knowledge about how children learn and grow and the appropriate pedagogical implications to maximize student learning. District salary schedules will typically rewards teachers for acquiring such advanced degrees, which undoubtedly encourages more teachers to pursue advanced studies. An important point to note is that districts rarely differentiate between the type of degree earned by a teacher or the quality of the institution from which it was attained when awarding bonuses for advanced studies. Another important policy feature relating to master's degrees is that some states even require advanced degrees for full certification (Hanushek & Rivkin, 2006). The most recent national data show the percent of full time public elementary and secondary school teachers with master's degrees is 44.5%. A minority of teachers has pursued even more advanced studies than this, with 6.4% of teachers in 2007/08 possessing an education specialist degree and 0.9% of teachers possessing a doctorate (Figure 3).

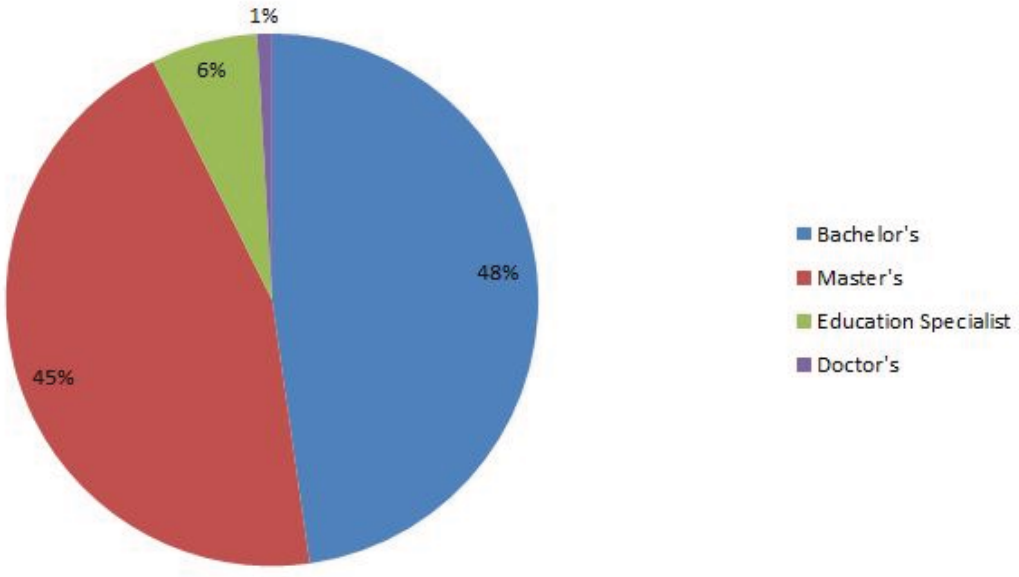

Figure 3. Percent of all full-time public elementary and secondary school teachers, by highest degree earned 2007/08. 'Education Specialist' includes certification of advanced graduate studies; Detail may not

sum to total because of rounding and omitted categories such as omission of teachers with less than a Bachelor's degree. U.S. Department of Education, National Center for Education Statistics, Schools and Staffing Survey (SASS), "Public Teacher Questionnaire," 2007-08.

Despite its surface plausibility and popular consensus that possession of an advanced degree leads to improved student achievement outcomes, subjective impressions are trumped by actual data on this topic. Hanushek (1997) conducted a comprehensive literature review of the research on master's degree impacts, weighing the results of thirty-four independent studies on the topic. Table 1 displays the results first by aggregating all studies on the topic then restricting the sample criteria to exclude low-quality estimates. Hanushek's table clearly demonstrates the remarkable finding that the vast majority of the research on the impacts of teacher education finds no relationship, either positive or negative, between attainment of a master's degree and teacher quality, as measured by student achievement outcomes. As a result, school leaders seeking to distinguish highly effective teachers during the hiring process and among teachers currently employed at their school cannot rely on graduate degree attainment as a single indicator of teacher quality.

Table 1

Percentage distribution of estimated effect of teacher education on student performance

Resources	Number of estimates	Statistically Significant: Positive	Statistically Significant: Negative	Statistically insignificant
All estimates				
Teacher education	170	9%	5%	86%
High- quality estimates				
Teacher education	34	0%	9%	91%

Note: This is a restricted version of a table produced by Hanushek and Rivkin (2006). High-quality estimates come from value-added estimation where the sample is drawn for individual students from a single state.

4.3 Certification or Licensure Status

Certification or licensure status is a measure of teacher qualifications, guaranteeing that a teacher has reached a pre-determined level of subject matter and pedagogical preparation and/or passed a series of exams as required by the state to ensure that no student is subjected to a teacher that fails to meet a minimal level of quality. Although the specifics of licensure vary by state, a standard teacher license is generally awarded to individuals who have satisfactorily completed a state-approved teacher preparation program, achieved the necessary college credits, and completed a student teaching experience. Additionally, most states require that candidates complete exams of subject matter and pedagogical knowledge such as the Praxis series. The purpose of teacher certification is to ensure that teacher candidates are adequately prepared for the profession. If the process of certification actually fulfills its purpose well, then one would expect that uncertified teachers would display lower levels of subject matter knowledge and inferior pedagogical skills. In other words, if teacher certification works how it is intended, we should be able to detect a discernible impact on student achievement.

A number of high profile reports such as the Carnegie Forum on Education and the Economy Task Force

on Teaching as Profession (*A Nation Prepared: Teachers for the 21st Century, 1986)* and the 1996 NCTAF report, *What Matters Most: Teaching for America's Future*, stress the importance of teacher content and pedagogical knowledge, recommending that all teachers graduate from state- approved teacher education programs. Calls such as this presuppose, of course, that conventional licensing actually increases teacher effectiveness; yet the research literature on teacher licensure is confusing upon initial inspection. On the one hand, distinguished Stanford professor and head of President Obama's transition team's working group on education policy, Darling-Hammond (2000) asserts that "measures of teacher preparation and certification are by far the strongest correlates of student achievement in reading and mathematics, both before and after controlling for student poverty and language status" (p. 1). On the other hand, noted economists Kane, Rockoff, and Staiger (2007) assure us that "certification doesn't guarantee a winner." A thorough examination of the evidence is required to distinguish fact from opinion.

The literature provides evidence of mixed impacts of teacher certification on teacher quality, as measured by student achievement outcomes. Wayne and Youngs (2003) describe the limitations of many reviews of studies on teacher certification. When interpreting results, for example, the reader should be aware of the strengths and weaknesses of various research designs: "The more trustworthy findings are those arising from studies that analyze individual students and their teachers and include a thorough set of controls for other potential determinants of student achievement." Similarly, readers should be wary of reviews that lump together test scores that measure different constructs from teachers at a variety of levels (secondary, elementary, etc.)

Kane et al. (2006) use six years of student-level data from grades three through eight in New York City public schools to evaluate the comparative effectiveness of certified, uncertified, and alternatively certified teachers. While they find large and persistent differences between teachers of the same certification status, they find only small impacts on student test performance between teachers of different certification statuses. When attempting to predict a teacher's future effectiveness, they find that the most reliable predictor is a teacher's classroom performance in the first two years on the job. Goldhaber and Brewer (2000) use Texas student achievement data to calculate value-added models of middle school student achievement on the Texas Assessment of Academic Skills (TAAS) from 1997/98 through 1998/99, finding that teachers with subject-matter certification in math perform better than other teachers but teachers with emergency certification credentials perform the same as teachers with standard certification.

Although teacher certification requirements exclude ill-prepared individuals from entering the classroom, overly restrictive certification policies may also discourage individuals who are actually very effective teachers from even applying (see for example Murnane, Singer, Willett, Kemple, & Olsen, 1991). Given that there is little evidence that teacher credentialing influences teacher effectiveness, the continued use of such policies in their current format requires scrutiny and debate.

4.4 Value- Added Measurements

Value-added measurement tools use complex statistical procedures that draw on multiple years of student achievement outcomes to isolate the contribution of individual schools and teachers to student growth (Buddin 2011; McCaffrey et al., 2003; Sanders & Horn, 1998). These relatively new tools have the potential to separate the effects of family background characteristics from school inputs to student achievement. Substantiation of differences in effectiveness between teachers can be informative for school leaders seeking to identify differences in experienced candidates at the recruitment stage and also when making assignment, development, and retention decisions regarding teachers currently working at their schools. While much work remains in the development of these tools regarding addressing measurement and statistical issues regarding the validity of these measures (Rothstein, 2009), the results uncovered here lend support to the notion that school leaders should begin discussions about the potential value of these new tools and start the process of reviewing and evaluating popular approaches to value-added measurements.

5 Conclusion

"It is difficult to identify those teachers who will prove more effective at the time of hire...better teachers can only be identified after some evidence on their actual job performance has accumulated" (Staiger & Rockoff, 2010).

Teachers appear to be the single most important school-based input affecting student achievement (Aaronson et al., 2007 ; Nye, Konstantopoulous, & Hedges, 2004; Kane et al., 2008; Rivkin et al., 2005 ; Kain, 2005; Rockoff, 2004; Sanders & Rivers, 1996). Educational outcomes depend more on the quality of the teacher a student is assigned to than on any other factor outside of the home, yet only a small proportion of the variation in teacher effectiveness is explained by the types of observable characteristics that are most commonly found in administrative data sets. In this review, I have examined the evidence on teacher experience, advanced degrees, and licensure or certification, finding small to null effects. Researchers' general lack of success in developing strongly predictive models based on these characteristics has resulted in continued debate about the causal relationship between specific teacher characteristics and student achievement growth. In response to the inconsistent, unpromising results of research on predictors of teacher quality, Winters (2012) suggests that principals and school leaders should instead be deriving a definition of teacher quality entirely from student learning. A teacher's independent contribution to student learning is the most meaningful measure of teacher quality and importantly, it cannot be consistently predicted based on observable characteristics, despite their often intuitive appeal.

In light of these findings, the implications for school leaders are as follows. The best strategy for attracting and raising teacher talent is to collect and link student-level achievement data to specific teachers and begin developing procedures that employ value-added modeling of student test scores as part of an overall evaluation of a teacher's effectiveness in the classroom. Results of such models can be used to target professional development opportunities where they will have the biggest impact, to identify high quality mentor teachers, and to target personnel policies towards recruiting and retaining the most effective teachers. The development of a quality teaching force is essential to improve educational outcomes for all students. School leaders need to seriously consider incorporating value-added measurement tools to judge recruitment, retention, development, and assignment decisions at the school level.

Future research is needed in two directions. The first is continued analysis and development of various econometric specifications for value-added models and descriptions of the degrees of uncertainty or imprecision associated with each. The second set of studies should analyze implementation of teacher evaluation tools that use value-added methodologies for teacher hiring, development, or retention decisions. Case studies of implementation efforts would be very informative for school leaders seeking to replicate such systems and wishing to anticipate potential challenges associated with implementation.

6 References

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics, 25*(1), 95-135.

Angrist, J. D., & Lavy, V. (2001). Does teacher training affect pupil learning? Evidence from matched comparisons in Jerusalem public schools. *Journal of Labor Economics, 19*(2), 343–369.

Buddin, R. (2011). *Measuring teacher and school effectiveness at improving student achievement in Los Angeles elementary schools* (Working Paper No. 31963). Retrieved from Munich Personal RePEc Archive website: http://mpra.ub.uni-muenchen.de/31963/

Clotfelter, C. T., Ladd, H. F., Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources, 41*(4), 778-820.

Darling-Hammond, L. (2000). Teacher quality and student achievement. *Educational Policy Analysis Archives, 8*(1).

Ferguson, R. (1998). Can schools narrow the black-white test score gap? In C. Jencks & M. Phillips (Eds.), *The black-white test score gap* (pp. 318-374). Washington DC: Brookings Institution Press

Goldhaber, D. D., & Brewer, D. J. (1997). Why don't school and teachers seem to matter? *Journal of*

Human Resources, 32(3), 505–523.
Goldhaber, D. D., & Brewer, D. J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis, 22*(2), 129–145.
Goldhaber, D. D., & Brewer, D. J. (2001). Evaluating the evidence on teacher certification: A rejoinder. *Educational Evaluation and Policy Analysis, 23*(1), 79–86.
Hanushek, E. A. (1992).The trade-off between child quantity and quality. *The Journal of Political Economy, 100*(1), 84-117.
Hanushek, E. A. (1996). A more complete picture of school resource policies. *Review of Educational Research,* 66, 397-409.
Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis, 19*(2), 141–164.
Hanushek, E. A. (2003). The failure of input-based schooling policies. *Economic Journal*, 113, 64-98.
Hanushek, E., Kain, J.F., O'Brien, D. & Rivkin, S.G. (2005). *The market for teacher quality* (Working Paper No. 11154). Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w11154
Hanushek, E.A. & Rivkin, S.G. (2006). Teacher quality. In E. Hanushek & F. Welsh (Eds.), *Handbook of the Economics of Education* (Vol. 2, pp. 1051 – 1078), Welsh, Amsterdam, NL: Elsevier B.V.
Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality, and student achievement. *Journal of Public Economics, 95*(7-8), 798-812.
Jacob, B. (2007). The challenges of staffing urban schools with effective teachers. *The Future of Children, 17*(1), 129-153.
Jacob, B., & Lefgren, L. (2004). Remedial education and student achievement: A regression-discontinuity analysis. *Review of Economics and Statistics, 86*(1), 226-244.
Jepsen, C., & Rivkin, S. (2002). *What is the tradeoff between smaller classes and teacher quality?* Retrieved from http://www.nber.org/papers/w9205.
Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2007). Photo finish: Certification doesn't guarantee a winner. *Education Next, 7*(1), 60-67.
Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review, 27*(6), 615–631.
Kane, T. J., & Staiger, D.O. (2005). *Using imperfect information to identify effective teachers.* Unpublished manuscript. Dartmouth College, Hanover, NH. Retrieved from https://www.dartmouth.edu/∼dstaiger/Papers/20
Kane, T. J., & Staiger, D.O. (2008). *Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation* (Working Paper No. 14607). Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w14607
Krei, M. (1998). Intensifying the barriers: The problem of inequitable teacher allocation in low-income urban schools. *Urban Education, 33*(1), 71-94.
Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis, 24*(1), 37-62.
McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2003). *Evaluating value-added models for teacher accountability.* Santa Monica, CA: The Rand Corporation.
Murnane, R. J., Singer, J. D., Willett, J. B., Kemple, J. J., & Olsen, R. J. (1991). *Who will teach? Policies that matter.* Cambridge, MA: Harvard University Press.
Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis, 26*(3), 237-257.
Rivers, J. & Sanders, W. (2002). Teacher quality and equity in educational opportunity: Findings and policy implications. In L.T. Izumi & E. M. Evers (Eds). *Teacher Quality* (pp. 13-24). Stanford, CA: Hoover Insstitution Press.
Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417-458.
Rockoff, J. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review, 94*(2), 247-252.

Rockoff, J.E., Jacob, B.A., Kane, T.J,. & Staiger, D.O. (2008). *Can you recognize an effective teacher when you recruit one?* (Working Paper No. 14485). Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w14485

Rothstein, J. (2009). Student sorting and bias in value added estimation: Selection on observables and unobservables, *Education Finance and Policy, 4*(4), 537-571.

Sanders, W. L., & Horn, S. P. (1998). Research findings from the Tennessee value-added assessment system (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education, 12*(3), 247-256.

Sanders, W. L., & Rivers, J. C. (1996). *Research progress report: Cumulative and residual effects of teachers on future student academic achievement: Tennessee value-added assessment system.* University of Tennessee Value-Added research and Assessment Center. Retrieved from http://www.cgp.upenn.edu/pdf/Sanders

Staiger, D. O., & Rockoff, J. E. (2010). Searching for Effective Teachers with Imperfect Information. *Journal of Economic Perspectives, 24*(3): 97–118.

Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*, 783-805.

Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*, 202-248.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73*(1), 89-122.

West, M., & Chingos, M. (2008, February). *Teacher effectiveness, mobility and attrition in Florida: A descriptive analysis.* Paper presented at Performance Incentives: Their Growing Impact on American K-12 Education Conference, Nashville, TN.

Whitehurst, G. J. (2012). The value of experiments in education. *Education Finance and Policy, 7*(2), 107-123.

Winters, M.A. (2012). *Teachers matter.* Rowman and Littlefield: Lanham, Maryland.

Woolfolk, A. E., & Hoy W.K. (1990). Prospective teachers' sense of efficacy and beliefs about control. *Journal of Educational Psychology*, 82, 81-91.

Woolfolk Hoy, A., Hoy, W. K., & Davis, H. (2009). Teachers' self-efficacy beliefs. In K. Wentzel & A. Wigfield (Eds.), *Handbook of motivation in school* (pp. 627-654). Mahwah, NJ: Erlbaum.

[2] http://www.cgp.upenn.edu/pdf/Sanders_Rivers-TVASS_teacher%20effects.pdf