Teacher Characteristics and Student Learning: A Comprehensive Assessment

Heather Hill
Harvard Graduate School of Education
Charalambos Y. Charalambous
University of Cyprus
Mark Chin
Harvard Graduate School of Education

The past decades have seen wide debates and scholarly inquiry about the teacher-level resources most helpful in promoting student learning outcomes. These resources tend to be organized into four broad categories, including experiences hypothesized to be important to teachers (e.g., post-secondary mathematics coursework, prior teaching, degrees, and certification type; see Clotfelter, Ladd, & Vigdor, 2007; Kane, Rockoff, & Staiger, 2008; Monk, 1994; Wayne & Youngs, 2003); teacher knowledge (e.g., mathematical knowledge for teaching, knowledge of students' ability and misconceptions; see Hill, Rowan, & Ball, 2005; Sadler, Sonnert, Coyle, Cook-Smith, & Miller, 2013); teacher mindset and habits (e.g., efficacy, locus of control, effort invested in teaching; see Muralidharan & Sundararaman, 2011; Philipp, 2007; Rose & Medway, 1981; Tschannen-Moran, Hoy, & Hoy, 1998); and professional supports for teaching (e.g., professional development). Together, analysts have suggested over a dozen ways to think about and measure the knowledge, training, and experiences teachers might use in helping students learn.

While many of the studies noted above show small positive associations between these characteristics and student outcomes as measured by standardized assessments, most studies tend to specialize in only one category, restricting the variables tested and thus potentially obscuring important relationships. Teacher experiences, for instance, are captured in district administrative or human resources databases and have been written about extensively in the economics of education literature, whose authors obtain access to such databases; papers in this field, however, tend not to include variables that must be directly measured. Studies of teachers' knowledge, on the other hand, collect large amounts of original data to measure that knowledge, a data collection burden so intense that authors rarely capture other key variables, such as mindsets and habits (e.g., Hill et al., 2005; Sadler et al., 2013). The same is true for studies of teacher efficacy, teacher professional development, and other key variables; in fact, seldom does a single study test more than one class of resource. Yet it seems likely that many of these factors are related; teachers with stronger post-secondary mathematical preparation, for instance, may also have stronger mathematical knowledge for teaching and may feel more capable of teaching the subject.

To untangle these relationships, we argue for a more comprehensive comparison of these resources, with the aim of understanding how they relate to one another and whether they individually and jointly contribute to student outcomes. To accomplish this aim, we draw on data from roughly 300 teachers participating in a study that measured the teacher resources named above as well as student outcomes on both a state standardized and a low-stakes, project-developed alternative mathematics assessment. We find that the correlations between most variables representing teacher resources are mild, and that variables from each of the four categories outlined above predict student outcomes. Specifically, teachers who take more mathematics content and methods courses and teachers whose knowledge—both mathematical and of student capabilities—is stronger have better student outcomes. Teacher-reported effort was also a significant predictor of state test performance, although other mindset variables were not. We discuss implications of these findings for hiring decisions and policy emphases.

Background

Below, we review evidence on each of our four categories of teacher background variables. We begin with our first category, teacher background and experiences.

Among the teacher background characteristics examined in education production function studies, teacher experience, teacher preparation and degrees obtained, as well as post-secondary mathematics coursework and certification type—all considered to function as proxies of teachers' capacity to teach—have been studied extensively. Among these variables, teacher experience was most often found to have a positive effect. Drawing on several quasi-experimental studies designed to explore the effect of teacher experience and student learning, Rice (2003) also reports that this effect has been positive, yet, more pronounced in the first years of teaching, especially for elementary school teachers. In her review, she concludes that a curvilinear effect could also be possible, given that student performance was found to increase with each year of teacher experience during the first three to five years of teaching, but then tapered off for more experienced teachers. By showing early-career teachers to typically have weaker student outcomes than their more experienced colleagues, more recent studies (Chetty, Friedman, & Rockoff, 2014; Kane et al., 2008) also empirically support that experience matters, at least for the first years of teachers' careers. Despite this generally positive result across studies, teacher experience did not account for much of the variance in student performance. For instance, on the basis of 18 studies investigating the magnitude of teacher effects, Nye, Konstantopoulos, and Hedges (2004) concluded that teacher experience could explain at most 5% of the variance in student performance.

Studies on teacher preparation and degrees obtained provide more mixed evidence of impact on teacher effectiveness (cf. Wayne & Youngs, 2003). One exception is that in mathematics, converging evidence suggests that high-school mathematics teachers who hold advanced degrees are more effective than those who do not (Goldhaber & Brewer, 2000; Rowan, Chiang, & Miller, 1997; Rowan, Correnti, & Miller, 2002; Wayne & Youngs, 2003). However, evidence regarding the impact of advanced degrees at the elementary level is mixed, ranging from non-significant effects (Wayne & Youngs, 2003) to even negative effects (Rowan et al., 2002).

The effects of teacher post-secondary mathematics coursework on student learning are again inconclusive. In one of the earliest meta-analyses examining studies conducted between 1960 and 1976, Begle (1979) reported that the number of post-secondary content courses – in this case, topics like college algebra – taken was positively associated with student achievement only in 10% of the examined studies and negatively associated in 8% of the analyzed studies. Similarly, mathematics methods courses, typically taught in schools of education with a focus on mathematics-specific pedagogy, were associated with positive effects in 24% of the cases and negative effects in 6% of the cases. In subsequent studies, mathematics content courses, typically taught in mathematics departments and focused on the content at the grade levels teachers will teach, were shown to have a curvilinear effect on student performance (Monk, 1994), with more pronounced effects at the high school level (Rice, 2003), and to be stronger for advanced, rather than remedial, students (Monk & King, 1994). Mathematics methods courses, on the other hand, were quite consistently found to have a positive effect on student learning (e.g., Monk, 1994; Rice, 2003).

Equally mixed and indeterminate are the results pertaining to teacher certification. Whereas high-school teachers holding a mathematics teacher certification were found to be more effective than those who do not (Rice, 2003; Wayne & Youngs, 2003), teacher certification was not shown to have a significant effect for elementary school teachers (Rowan et al., 2002; Hill et al., 2005). This picture becomes even more inconclusive as data from different grades and different subject matters are employed. For example, drawing on six years of panel data in mathematics and reading from fourth- through eighth-grade students and their teachers in New York City, Kane, Rockoff, and Staiger (2008) found little difference in the academic achievement impacts of certified, uncertified, and alternatively certified teachers. However, when comparing the effect of teachers with very weak credentials (e.g., low licensure test scores, a lateral entry license, and a certification not in the specific subject) to those with strong credentials (e.g., National Board Certification, degree from a very competitive undergraduate institution), the effect of teacher credentials becomes more prominent (cf. Clotfelter et al., 2007).

Our second category, teacher knowledge, ranks high among the desired teacher characteristics, both in written policy documents (Council of Chief School Officers, 2011; National Board for Professional Teaching Standards, 1989) and policy itself, in that most states have teacher certification tests with minimum bars for entry into the profession (e.g., ETS' *PRAXIS*). Teacher knowledge is most often thought of as multi-faceted, containing components that reflect the various tasks they engage in classrooms, such as representing content to students, designing tasks, and understanding student thinking. Shulman (1986; 1987) categorized these broadly into content knowledge, general pedagogical knowledge, pedagogical content knowledge, and knowledge of learners and their characteristics, among other topics. In mathematics, the topic of this study, Ball, Thames, and Phelps (2008) elaborated this list, arguing that teachers have "specialized content knowledge" that goes beyond basic knowledge of mathematics, and for the presence of two distinct sub-domains within pedagogical content knowledge: knowledge of content and students, and knowledge of content and teaching. Other categorizations also exist (e.g., Davis & Renert, 2013; Rowland, Huckstep, &Thwaites, 2005), although a recent review (Depaepe, Verschaffel, & Kelchtermans, 2013) suggests considerable consensus around the main categories.

A number of studies have empirically linked specific aspects of teacher knowledge to student outcomes. Some focus around teachers' knowledge of the mathematics taught to students, finding an association between teachers' pure content knowledge and students' test score outcomes (Harbison & Hanushek, 1992; Metzler & Woessmann, 2012; Mullens, Murnane, & Willett, 1996; Rowan et al., 1997). Others focus on measuring Ball et al.'s category "mathematical knowledge for teaching," finding that teachers with stronger command of areas such as explanations for mathematical ideas and procedures, alternative solution methods, and ways to visually model the content taught students who performed better on both low-stakes tests and state assessments (Hill et al., 2005; Hill, Kapitula, & Umland, 2011; Rockoff, Jacob, Kane, & Staiger, 2011). Still others focus around knowledge of students in other forms, including the accuracy with which teachers can predict their current students' performance (Carpenter, Fennema, Peterson, & Carey, 1988; Helmke & Schrader 1987) and anticipate or interpret common student misconceptions (Sadler et al., 2013). While additional types of teacher knowledge are hypothesized to exist, they have not yet been operationalized on the scale necessary to detect relationships between teacher knowledge and student outcomes.

A third category of teacher attributes thought to contribute to student learning concerns teacher mindset and habits. Under this umbrella, scholars have examined attributes such as teachers' efficacy beliefs, locus of control, and effort invested in teaching (e.g., Muralidharan & Sundararaman, 2011; Lavy, 2009; Philipp, 2007; Rose & Medway, 1981; Tschannen-Moran et al., 1998). Among these constructs, teacher efficacy beliefs have been explored more systematically. Defined as teachers' sense of ability to organize and execute teaching that promotes learning (Bandura, 1997; Charalambous, Philippou, & Kyriakides, 2008; Usher & Pajares, 2008), teacher efficacy beliefs pertain to a future-oriented judgment capturing teachers' *perceptions* about their competence rather than their actual efficacy (Hoy & Spero, 2005). Teacher efficacy beliefs in mathematics and other subject matters have consistently been found to positively relate to teachers' behavior in the classroom and the quality of their instruction (Graham, Harris, Fink, & MacArthur, 2001; Justice, Mashburn, Hamre, & Pianta, 2008; Stipek, 2012; Tschannen-Moran et al., 1998; Tschannen-Moran & Hoy, 2001; Woolfolk, Hoy, Hoy, & Davis, 2009). More importantly, they have been shown to be predictive of student learning outcomes, both cognitive (Ashton & Webb, 1986; Guo, Piasta, Justice, & Kaderavek, 2010; Raudenbush, Rowan, & Cheong, 1992; Ross, 1992) and affective (Anderson, Greene, & Loewen, 1988; Soodak & Podell, 1996).

Teacher locus of control pertains to the attributions that teachers make about students' successes and failures. In essence, this construct taps the extent to which teachers feel they can influence their students' outcomes or if, alternatively, they believe that those outcomes mostly hinge on non-classroom and school factors (e.g., students' socioeconomic background and parental support) (cf. Rose and Medway, 1981). Inspired by Rotter's (1966) work on internal versus external control of reinforcement, Rand researchers (Armor et al., 1976) showed the two items employed to capture teacher locus of control to significantly relate to teachers' success in teaching reading to minority students. The positive effect of teacher locus of control was subsequently documented in other studies (Berman & McLaughlin, 1977; Rose & Medway, 1981).

Teacher effort has been studied less extensively than the other two constructs pertaining to teachers' mindset and habits. Using a quasi-experimental design, Lavy (2009) observed that a positive effect of teacher merit pay on students' outcomes was mediated by increased teacher effort, particular in the area of after-school tutoring.

Our fourth category comprises resources teachers may access from their immediate environment, including professional development and school-specific resources.[1] Although promising results have resulted from specific professional development programs (e.g., Carpenter, Fennema, Peterson, Chiang, & Loef, 1989), large-scale studies that use surveys to measure teachers' professional development experiences have been less optimistic (Harris & Sass, 2011; Jacob & Lefgren, 2004), suggesting that while specific programs may work, the variety more typically available to teachers may have little effect. Effectiveness may also differ by the format of professional development offering (e.g., individual coaching versus group-based learning experiences; for a successful coaching program, see Campbell & Malkus, 2011), or by particular focus—for instance on student thinking, mathematical content, general pedagogy, or a combination. Teachers may also be able to glean concrete resources, such as materials and manipulatives, and psychological resources, such as the respect of others and pleasant working conditions, from their school.

In sum, within each of the four categories examined, different variables have been found to significantly relate to student outcomes. However, studies that bring together variables from different categories—to untangle the individual and joint effect of these variables on student performance and learning—are scarce. It is toward this direction that the present study aimed to contribute.

Data and Methods

[1] Here we categorize professional development opportunities with the other measures related to resources accessible to teachers, which reflects how it typically is considered to affect teacher experiences. In our analyses, we consider uptake of professional development first as a mindset and belief, reflecting teachers' predilection towards using such opportunities, and second as a resource.

The data we use in our analyses come from the National Center for Teacher Effectiveness (NCTE) main study, which spanned the academic years from 2010-11 to 2012-13. The study, which developed and aimed to validate several measures of mathematics teacher effectiveness, collected data from fourth- and fifth-grade teachers and their students in four, large urban East coast public school districts. At the time of recruitment, these districts were revising their teacher evaluation systems, but had not yet fully implemented changes. Furthermore, principals of the schools recruited for NCTE agreed to allow the study to randomly assign classroom rosters to participating teachers in 2012-13. As such, schools were also required to have at least two fourth- or fifth-grade teachers instructing self-contained, non-specialized (i.e., special education) classrooms of students. NCTE recruited 583 teachers across the four districts, of which 328 matriculated into the study.

Sample

This analysis uses the data of 306 teachers and the 10,233 students in their classrooms over the three study years. Twenty-two teachers were excluded because their students and classrooms failed to meet restrictions imposed on our models predicting student test performance, described in more detail below.

Among all teachers in our sample, 11% reported themselves to be a novice teacher at entry into our study. A relatively high proportion of the sample is traditionally certified (86%), and roughly half have a bachelor's degree in education (52%). A small proportion has a math-specific elementary certification (14%), and a relatively large fraction reported possessing a master's degree (77%). Student demographics reflected the urban settings in which we conducted this study, with 64% of students free- or reduced-price lunch eligible (FRPL), 10% qualifying for special education, and 20% designated as English Language Learners (ELL) at the time of the study. Moreover, a notable percentage of the participants were either Black (40%) or Hispanic (23%).

Data Sources and Reduction

Data collection relied upon several instruments including, for the purposes of this analysis, a background and experience questionnaire administered once to teachers in their first year of participation with NCTE; a fall questionnaire, administered each school year and comprising questions measuring teachers' mathematical knowledge as well as questions related to teachers' mindsets, beliefs, and perceptions; a spring questionnaire, administered each school year and comprising items assessing teachers' knowledge of students; student performance on a project-developed mathematics assessment (see Hickman, Fu, & Hill, 2012); and district administrative data, which provided information on students' state standardized test performance in addition to student demographic information.

Below we describe the specific measures we consider when connecting teacher characteristics to student learning. These measures, which we break into the four overarching categories described in our literature review (i.e., *preparation*, *knowledge, mindset/belief*, and *resource* measures), were selected and grouped based on prior theoretical and empirical evidence supporting their importance for student learning. When constructing estimates of teacher performance on these constructs, we de-mean scores to account for the sometimes sizeable differences in district-aggregated teacher characteristics, such as average teacher knowledge scores. We also do so because student state standardized test performance, one of our outcomes, is also standardized within district to account for differences in tests. After district de-meaning, teacher characteristic scores used in models predicting student test performance is standardized to have a mean of zero and a standard deviation of one for the sample of teachers.

Teacher experience measures. We used teacher responses to eight different items to develop the following seven measures:

- A dichotomous variable indicating a novice teacher (i.e., two or fewer years of reported experience);
- A categorical variable indicating teachers' reported number of undergraduate or graduate-level classes covering college-level mathematics topics (*math courses*). Responses to this variable range from 1 ("no classes") to 4 ("six or more classes");
- A categorical variable indicating teachers' reported number of undergraduate or graduate level classes focused on mathematics content for teachers (*math content courses*). Responses to this variable range from 1 ("no classes") to 4 ("six or more classes");
- A categorical variable indicating teachers' reported number of undergraduate or graduate level classes focused on methods for teaching mathematics (*math methods courses*). Responses to this variable range from 1 ("no classes") to 4 ("six or more classes");
- A dichotomous variable indicating a teacher who reported completing a traditional teacher education program prior to taking his or her first teaching job (*traditionally certified*), as opposed to participating in an alternative certification program (e.g., Teach for America) or not having any formal training. 81% of NCTE teachers reported being traditionally certified; we collapse the other two certification categories due to smaller samples (i.e., 7% of teachers reporting either);
- A dichotomous variable indicating a teacher's possession of a bachelor's degree in education (mean = 0.524);
- A dichotomous variable indicating a teacher's possession of a certificate in the teaching of elementary mathematics (mean = 0.147); and,
- A dichotomous variable indicating possession of any master's degree (mean = 0.771)

Knowledge measures. From the fall and spring teacher questionnaires, we analyze three measures of teacher knowledge: *MKT/MTEL*, capturing performance on items assessing teachers' Mathematical Knowledge for Teaching (Hill et al., 2005) and items from the Massachusetts Test of Education Licensure; teachers' *accuracy* in predicting student performance on items from the project-developed test; and teachers' knowledge of students' mathematical misconceptions (*KOSM*). The *MKT/MTEL* measure possessed a marginal reliability of 0.92 and the adjusted intraclass correlations of the teacher *accuracy* scores ranged from 0.71 to 0.79.[2] The *KOSM* measure appeared less able to differentiate teachers, reporting marginal reliabilities of 0.21 and 0.40 of scores for fourth and fifth grade teachers, respectively.

[2] Intraclass correlations were adjusted for the modal number of *accuracy* items teachers responded to. This adjustment provides an estimate of reliability that is more reflective of measure scores, which incorporate teacher responses to several items as opposed to a single item.

For more information on the construction and validity of these knowledge measures, please see Bacher-Hicks and colleagues (2015) and Hill, Chin, and Blazar (in preparation).

Mindset/Belief measures. We use teacher responses from items on the fall questionnaire to estimate four scores reflecting the following constructs of teachers' mindsets or beliefs: effort, efficacy, locus of control, and the amount of time spent participating in professional development (*Teacher PD*). The last requires some explanation. Professional development may lead directly to increases in teacher productivity; in a given year, for instance, participation in this activity may lead to subsequent increases in students' test scores; we examine this possibility below, using a model that accounts for year-to-year variability in teachers' reported attendance at professional development opportunities. Here, we include teachers' average professional development across multiple years as a measure of teachers' efforts toward improvement.

[Insert Table 1]

Table 1 above shows the items used for each measure, the item sources, and the internal consistencies of composites. Because the NCTE study asked teachers these questions across several administrations of surveys, we leveraged the additional information gained from multiple years of data when estimating scores for each *mindset/belief* measure using the following equation:

$$(1)\ \mathrm{TQ}_{yt} = \beta_0 + \alpha_y + \mu_t + \varepsilon_{yt}$$

The outcome in equation (1), TQ_{yt}, represents the average of teacher *t*'s responses, within year *y*, across the items of each respective construct. The model controls for differences in response level across years using year fixed effects, α_y. μ_t is a random effect for teacher *t*, the best linear unbiased predictor, and also captures each teachers' score on effort, efficacy, locus of control, or teacher PD.

Resource measures. Teachers also respond to nine questions on the fall questionnaire that probe their perceptions of the resources provided at the schools in which they instruct (*school resources*); the items and internal consistency estimates can also be found in Table 1. We estimate teacher scores for this measure using equation (1).

Although we consider teachers' teacher PD scores, when estimated from responses to survey items across years, as a *mindset/belief* measure in our primary models, teacher uptake of professional development opportunities may also vary from year to year. In an effort to isolate the effect of a given year's PD, we average responses to teacher PD items within a year and use these estimates in additional analyses alongside the school resource variable, similarly estimated within a year.

Imputation of missing data. Due to changes to the items asked in teacher questionnaires across survey administrations and changes to the population of teachers participating in the NCTE study across academic years, not every teacher in our sample of 306 has scores for each measure listed above. 197 of the 326 total teacher-grade-level observations had no missing data and 88% of such combinations missed scores for at most one measure. For cases of missing data, we imputed scores for teachers using the district mean; for categorical or dichotomous variables, teachers received district modal values.[3] In our models predicting student test performance, we include dichotomous variables indicating whether a teacher was missing data from a specific source (i.e., dichotomous variables for missing data from the background questionnaire, the fall questionnaire, or the spring questionnaire). In sensitivity analyses, described in the results section, we explore how our results describing the relationship between teacher characteristics and student test performance change when we make different choices about how to handle missing data.

Analysis Strategy

We use two primary strategies in our analyses. First, we correlate teacher scores from all the measures listed above. Correlational analyses allow us to better understand the characteristics of teachers in our sample and their relationship to one another. Furthermore, because several of the variables theoretically relate to one another, observed empirical patterns of strong relationships signal the possibility for multicollinearity influencing regression results.

Second, we predict student test performance on both state standardized mathematics tests and the project-developed mathematics test using the following multilevel model, which nests students within teacher-year combinations, which are subsequently nested within teacher:

$$(2)\ y_{spcgyt} = \beta_0 + \alpha X_{sy-1} + \delta D_{sy} + \phi P_{pcgyt} + \kappa C_{cgyt} + \eta + \omega \theta_t + \mu_t + \nu_{yt} + \varepsilon_{spcgyt}$$

The outcome, y_{spcgyt}, represents the test performance on either the state standardized or project-developed mathematics test of student *s*, in classroom *p*, in cohort (i.e., school, year, and grade) *c*, taking the test for grade *g*,[4] in year *y*, taught by teacher *t*. Equation (2) contains the following controls:

- X_{sy-1}, a vector of controls for student prior test performance;
- D_{sy}, a vector of controls for student demographic information (i.e., race or ethnicity, gender, FRPL eligibility; special education status; and ELL status);
- P_{pcgyt}, classroom-level averages of X_{sy-1} and D_{sy} to capture the effects of a student's peers;
- C_{cgyt}, cohort-level averages of X_{sy-1} and D_{sy} to capture the effect of a student's cohort;
- η, school, district, and grade-by-year fixed effects;
- θ_t, a vector of teacher-level scores for different sets of characteristic measures;[5]
- μ_t, a random effect on test performance for being taught by teacher t; and,
- ν_{yt}, a random effect on test performance for being taught by teacher t in year y.

[3] In one district, two categories tied for the most frequent response for the math courses variable. In this case, teachers received the whole sample modal value.
[4] Fewer than 1% of students took the state standardized mathematics test at a different grade level than the project-developed mathematics test. We consider this percentage to be negligible for the purposes of our analyses.
[5] Teacher *accuracy* and *KOSM* scores are included in the model at the teacher-grade level, and the indicator for being a novice teacher is considered at the teacher-year level.

The model represented by equation (2) contains controls that are used by many states and districts when estimating value-added scores for teachers. Furthermore, to ensure that our model predicts the test performance of typical students and classrooms, we impose the following restrictions to arrive at our final aforementioned sample of 10,233 students: students missing any model controls are excluded; students who skipped or repeated tests, based on the grade-level for either the state or project-developed assessment, are excluded; classrooms where greater than 50% of students do not have baseline scores for either outcome measure are excluded; classrooms where greater than 50% of students are considered special education students are excluded; and classrooms with fewer than five students are excluded.

We recover our primary parameters of interest, the relationship between different teacher characteristics to student test performance, from ω. We estimate equation (2) for each student outcome measure multiple times, using a different set of teacher characteristic variables, θ_t, across each iteration; specifically, we estimate equation (2) when including: only *teacher experiences* measures; only *knowledge* measures; only *mindset/belief* measures; only *resource* measures; and all measures.[6] When running models using different independent variable controls and different student outcomes, we hope to find consistent patterns of the magnitude and statistical significance of effects within measures. Finding such evidence would lend additional support to the importance of specific teacher characteristics. Finally, we investigate the joint importance for student test performance of teacher characteristics within each category using two different estimates. First, we run a Wald test to test the joint significance of these variables in predicting outcomes of all measures within each category; this test can offer support, agnostic to potential issues of multicollinearity, for the importance of overarching categories of teacher characteristics for learning. Second, we estimate how much variance in teacher effects on student achievement each group of characteristics explains (an "adjusted pseudo R-squared"; see Bacher-Hicks et al., 2015). Prior research has largely found that differences between teachers, in terms of effects on their students' performance on tests, to be largely unexplained by observable characteristics of teachers (see Palardy & Rumberger, 2008). Estimation of the adjusted pseudo R-squared statistic of combinations of measures thus potentially allows us to corroborate prior research, and helps us describe the amount of variance in teacher quality still left unexplored by the several measures theorized to comprise a large fraction of what makes teachers effective.

Results

We start by discussing the correlations among the variables explored. Table 2 below shows that few notable correlations arise between our independent variables.

[Insert Table 2]

Teachers' reports of completing math content and math methods courses were highly correlated (*r*=0.77). Traditionally certified teachers were more likely to possess a bachelors' degree in education (*r*=0.59), consistent with the conventional training and certification processes in most states. Overall, novice teachers in our samples took fewer courses, did not possess master's degrees, and felt less efficacious; these patterns reflect what we would expect

[6] As noted in the data sources and reduction section, we also estimate equation (2) including only the *resources* variables, considering the within-year teacher PD score as a resource variable in addition to the school resource variable.

intuitively, as newer teachers likely have had fewer opportunities to attain these additional developmental experiences or practice different teaching behaviors. Teachers who reported receiving their bachelor's degree in education were less likely to also have completed a master's degree (r= -0.34). There was a notable relationship between reported mathematics content courses with reports of more time spent on grading papers and preparing for class (*effort*) (r=0.25); effort also related to reports of taking more professional development (r= 0.33), perhaps a sign that a particular group of teachers in our sample had more time or inclination to invest in their work. Notably, attendance at math content or methods courses was not associated with the three teacher knowledge measures.

Looking further into just measures from the *knowledge, mindset/beliefs,* and *resources* categories, we further observe few patterns. Surprisingly, the two measures of teachers' knowledge of students did not correlate very strongly with one another (r= 0.08), though teachers' abilities to predict their students' mathematical capabilities (*accuracy*) were found to be related with teachers' knowledge of mathematics, general and specific for teaching (*MKT/MTEL*; r= 0.25). Teacher efficacy was also negatively related to locus of control (r= -0.27), and was positively related to the perceived resources provided by teachers' schools (r= 0.25). The former observed relationship reflects what might be expected intuitively: teachers feel they are more capable of implementing effective teaching practices when they feel as though student learning can be influenced. The latter finding suggests that teachers' sense of efficacy is closely tied to the available resources of their schools, and complements prior work focusing on teacher efficacy as a school-level measure. However, with the exception of math content and methods courses, none of these correlations was strong enough to warrant concern about multicollinearity. For analyses below, we combined math content and methods courses by averaging teachers' reports of the two.

[Insert Table 3]

Table 3 shows regressions predicting student performance on the NCTE and state test outcomes. Among self-reported teacher experiences, only the completion of math content/methods courses was significant in both the project and state tests, with a coefficient size totaling roughly 4 to 5% of a student-level standard deviation. And although none of the certification pathways were themselves significant, possession of a bachelor's degree in education was also marginally positively associated with student outcomes on the state tests. As noted above, possession of an education bachelor's degree was more often observed in traditionally certified teachers. Finally, traditional certification is associated with a negative relationship to student outcomes on the project test. Only the math methods/content course reports, however, remained significant at the $p < 0.05$ level in the final models. Beyond these variables, no other variables within this category appeared significantly related to student outcomes, including experience and elementary mathematics certification. Noticeably, the overall variance explained by these variables is quite low, and the Wald test indicated that the measures together were not jointly significant for predicting performance on the project-administered assessment, and were only marginally significant (p=0.06) for the state test.

In the knowledge category, teachers' *accuracy* scores predicted their students' outcomes on the project-administered assessment as well as the state assessment. *MKT/MTEL* was a marginally significant predictor of student outcomes in the final model of the project-administered assessment (p<0.10). Our measure of knowledge of student misconceptions, the least reliable of the three, showed no relationship to student outcomes. These results suggest that two of the three aspects of teacher knowledge evaluated here may possess some positive relationships with student outcomes; this is also apparent from the relatively large amount of variance explained by this category (for the project-developed test, 15%) and the strong significance of the Wald test for both outcomes. These results also suggest that teaching-related mathematical knowledge and predictive accuracy, though correlated with one another, may be individually important – and thus separate – contributors to student growth.

In the mindset and habits category, neither the self-efficacy nor the locus of control measures predicted performance on the state or project assessment in the final analysis. However, locus of control – which measured teachers' beliefs on how much factors outside their control (i.e., student effort, intelligence, or family background) influence student learning – was negatively related to student outcomes on the project-administered test when included with just the other personal resources variables. Professional development, conceptualized in these models not as a causal effect of professional development but as a marker of effortful improvement, similarly showed no relationship to student outcomes. By contrast, teachers' self-reported effort – the number of hours spent grading, preparing for class, and tutoring students outside of regular school hours – predicts performance on the state test. To check the intuition that the last may have been a driver of this result, perhaps as teachers tutor students in preparation for state assessments, we removed the tutoring item from the scale and found the same result (when using all teacher characteristics as controls, $b=0.051$, $p < 0.01$). Despite this striking finding, the amount of variance explained by this category is small (6%) and the Wald test barely significant ($p=0.04$).

Finally, professional supports for teaching did not perform as expected by theory and some prior literature. Teachers' reports of school resources had no relationship to either the state or project-administered assessment outcomes in the cross-sectional models (Table 3), and a negative relationship when year-specific measures were used to predict year-specific outcomes on the state assessment (see Table 4). Professional development conceptualized as a year-specific resource similarly shows a null result.

[Insert Table 4]

We now turn to the sensitivity analysis run to explore the robustness of our findings. As noted above, we utilize a sample of 306 teachers and all their students in our analyses predicting student outcomes using teacher characteristics, despite many teachers in our sample missing data on one or more measures. Thus, we explore how observed relationships change from our main analysis, which uses imputed scores for missing data, under different samples of teachers and students. Consistent findings across samples would support the conclusions we make from the larger sample of teachers and students.

When including in models all the teachers that have non-imputed scores for all measures within groupings (i.e., *teacher experiences, knowledge, mindsets/beliefs, resources*, and *all*), our findings largely stay the same. In predicting student test performance on the project-developed assessment, math content/methods courses and teacher *accuracy* predict outcomes across models. Findings associated with being traditionally certified are attenuated under this sample. In predicting performance on state standardized tests, teacher effort and enrollment into math content/methods courses continues to strongly predict student outcomes, though the significance of the point estimate for teacher *accuracy* diminishes despite maintaining a similar magnitude.

For both outcomes, teacher self-efficacy and enrollment into math courses, show stronger, more significant relationships with outcomes when compared to the relationships observed using the sample of teachers with imputed data; for efficacy, this relationship is positive, and for math courses, this relationship is negative. We see these same patterns when including in models only those teachers that have data for all characteristics ($n = 190$). These findings corroborate our initial results of the importance of math content/methods courses, teacher *accuracy*, and effort for student test performance, but also suggest that our imputation method for missing data may mask the effect of teacher efficacy and math courses, as well.

Discussion

We initiated this exploration noting that although scholarly interest has been directed toward exploring the effect of several teacher-level characteristics thought to contribute to student learning, in most of the cases, these characteristics were studied in isolation. By bringing together characteristics from four main categories—teacher experiences, teacher knowledge, teacher mindsets and habits, and reported professional support for teaching—we aimed at investigating their joint effect in explaining student learning. We undertook this investigation by considering two outcome variables: student learning as captured on both a state standardized and project-developed alternative mathematics assessment.

By and large, results from both student-level outcomes under consideration pointed to the same characteristics as potentially important. Consistent with prior literature (Monk, 1994; Rice, 2003), we found that the completion of math content/methods courses had a positive effect on student learning. This is remarkable in an era in which many teacher preparation programs – particularly alternative entry pathways – do not feature content-specific teaching coursework. It suggests that such coursework may be an important support for elementary teachers; however, selection effects (e.g., teachers more comfortable with mathematics enroll in more such courses) cannot be ruled out.

Similarly, in line with recent research findings on teacher knowledge, two of the teacher knowledge measures we employed –teacher *accuracy* and *MKT/MTEL*—were found to not only have a positive effects, but to explain more variance in student outcomes than the other categories examined. These teacher knowledge variables are by and large uncorrelated with teachers' mathematics methods and content coursework, suggesting that teachers arrived at more such knowledge through other means.

Contrary to what other prior studies have reported (cf. Carpenter et al., 1989; Chetty et al., 2014; Tschannen-Moran & Hoy, 2001) teacher self-efficacy beliefs and professional development were not found to have any effect on student learning for either of the two outcomes examined. Teacher experience effect sizes are similar to other recent reports (e.g., Kane et al., 2008) but marginally related to student outcomes on the state assessment, perhaps a result of only a small number of our teachers (roughly 30) holding this designation. Other reasons may explain the lack of relationship between self-efficacy and professional development. For example, our survey did not allow delving deeper into the quality of the professional development the participants received; investigation into one NCTE district, which strongly aligned instructional goals and teacher resources (i.e., coaching and professional development) to components emphasized by the state standardized test, for example, revealed significant correlations between hours of professional development attended with student outcomes (personal communications). Regardless of what these explanations might be, these findings seem to suggest that attributes that have traditionally been reported to have significant effects on student learning may not be that influential, when other teacher-level resources are brought to the equation.

At the same time, our findings suggest that other resources which have not been given particular attention in the literature so far might be worth some closer consideration. This was particularly the case with teacher effort, which had a positive effect on student learning as measured by state test results. Because we cannot make a causal attribution in this study, this result calls for further investigation. If replicated in other studies with different teacher populations and improved designs, this finding might suggest that, as in other professions, effort can help increase effectiveness, and consequently (student) outcomes. At the same time, it would point to future research directions, suggesting that more complex interactions than those examined in our work be considered: instead of simply exploring the additive effect of different teacher-level attributes—as we did in our work—interactions between and among them could also be considered. For example, how do teacher knowledge and effort interact in informing student learning? Does the effect of one of them mediate the effect of the other? If so, in what particular ways?

In larger perspective, these findings suggest that despite some consistent patterns, there does not seem to be one "silver bullet" teacher characteristic that explains teacher effectiveness. Even the variables found to have a significant contribution were found to have small effects; additionally, they explain a moderate, at best, percentage of the variance in student learning. This result has several implications. To start, it calls for future studies that expand their lens and consider multiple teacher characteristics: only by doing so, will we be able, as a research community, to really weigh the contribution of different teacher attributes. While in isolation the effect of different variables might be significant, these significant contributions might cease to hold when other teacher attributes are brought to the fore. Conversely, and provided that more complex interactive effects are explored, variables that seem to not contribute significantly to student learning, might turn out to do so, when seen through a meditator lens. At the same time, this finding suggests that focusing solely on teacher attributes—as reported in surveys—might not be sufficient to understand how students learn. Instead, future attempts should be directed at exploring how teachers deploy these resources and capitalize on them during instruction. Additionally, apart from these teacher attributes, scholarly interest can be directed to also considering teacher behaviors during instruction.

We conclude by considering the implications of our findings for hiring decisions and policy emphasis. Although we considered a conglomerate of teacher attributes—much larger than typically examined in prior studies—we were not able to identify even a single teacher characteristic that had a strong contribution to student learning; nor were we able to explain a notable portion of the variance in student learning. Collectively, these findings challenge the common approach often pursued when hiring teachers to search for particular teacher attributes among the pool of candidates. Even though these attributes could play some role in teacher effectiveness, they definitely seem to tell only a portion of the story—admittedly, a small one. Hence, hiring decisions should not be based solely on teacher attributes like those examined in our study. Candidates' performance while teaching in real settings or the instructional decisions they make in virtual environments need to be also weighted when making appointment decisions.

At the same time, the fact that we did not find very strong correlations between variables often assumed to be theoretically connected suggests that these correlations need to be more carefully scrutinized—both at theoretical and empirical levels. For example, mathematics methods courses are often assumed to lead to better knowledge, especially knowledge for teaching the subject matter. Although several explanations might be offered for the lack of significant correlations between these two traits in our work, we cannot overrule the possibility that indeed mathematics methods courses do not necessarily strengthen teachers' content knowledge—at least in a straightforward manner as implied when considering linear associations between these traits. To the extent that this hypothesis holds, it underlines the importance of not only examining if such courses contribute to teachers' learning, but foremost exploring the mechanisms through which such a contribution could be feasible. Studies that are currently undertaken toward this direction (e.g., Steele, Hillen, & Smith, 2013) show the promise of this line of work.

Research during the past three decades has helped emphasize the critical role that teachers have for student learning. Yet determining what makes an effective teacher, in general, and how certain teacher attributes support teacher effectiveness, in particular, remains an open issue. Therefore, in the years to come scholarly interest should move from simply exploring *what* teacher attributes contribute to student learning to *how* these attributes, individually and interactively, make a difference to teacher effectiveness, and consequently to student learning.

Works Cited

Anderson, R. N., Greene, M. L., & Loewen, P. S. (1988). Relationships among teachers' and students' thinking skills, sense of efficacy, and student achievement. *Alberta Journal of Educational Research*.

Armor, D. J., Conry-Oseguera, P., Cox, M., King, N. J., McDonnell, L. M., Pascal, A. H., ... & Zellman, G. L. (1976). Analysis of the School Preferred Reading Program in Selected Los Angeles Minority Schools.

Ashton, P. T., & Webb, R. B. (1986). *Making a difference: Teachers' sense of efficacy and student achievement*. Longman Publishing Group.

Bacher-Hicks, A., Chin, M., Hill, H. C., & Staiger, D. O. (2015). *Explaining teacher effects on achievement using measures from multiple research traditions.* Manuscript submitted for publication.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching what makes it special?. *Journal of teacher education*, *59*(5), 389-407.

Bandura, A. (1997). Self-efficacy: The exercise of control.

Begle, E. G. (1979). Critical Variables in Mathematics Education: Findings from a Survey of the Empirical Literature.

Berman, P., & McLaughlin, M. W. (1977). *Federal Programs Supporting Educational Change, Volume VII: Factors Affecting Implementation and Continuation*. Santa Monica, CA: The RAND Corporation.

Campbell, P. F., & Malkus, N. N. (2011). The impact of elementary mathematics coaches on student achievement. *The Elementary School Journal*, *111*(3), 430-454.

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C. P., & Loef, M. (1989). Using knowledge of children’s mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal*,*26*(4), 499-531.

Carpenter, T. P., Fennema, E., Peterson, P. L., & Carey, D. A. (1988). Teachers’ pedagogical content knowledge of students’ problem solving in elementary arithmetic. *Journal for Research in Mathematics Education, 19*(5), 385–401.

Charalambous, C. Y., Philippou, G. N., & Kyriakides, L. (2008). Tracing the development of preservice teachers’ efficacy beliefs in teaching mathematics during fieldwork. *Educational Studies in Mathematics*, *67*(2), 125-142.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood.*American Economic Review*, *104*(9), 2633-79.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects.*Economics of Education Review*, *26*(6), 673-682.

Council of Chief School Officers. (2011). *InTASC Model Core Teaching Standards*. Retrieved from http://www.ccsso.org/resources/publications/intasc_model_core_teaching_standards_2011_ms_word_version.html

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education policy analysis archives*, *8*, 1.

Davis, B., & Renert, M. (2013). Profound understanding of emergent mathematics: broadening the construct of teachers' disciplinary knowledge.*Educational Studies in Mathematics*, *82*(2), 245-265.

Depaepe, F., Verschaffel, L., & Kelchtermans, G. (2013). Pedagogical content knowledge: A systematic review of the way in which the concept has pervaded mathematics educational research. *Teaching and Teacher Education*, *34*, 12-25.

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.

Graham, S., Harris, K. R., Fink, B., & MacArthur, C. A. (2001). Teacher efficacy in writing: A construct validation with primary grade teachers. *Scientific Studies of Reading*, *5*(2), 177-202.

Guo, Y., Piasta, S. B., Justice, L. M., & Kaderavek, J. N. (2010). Relations among preschool teachers' self-efficacy, classroom quality, and children's language and literacy gains. *Teaching and Teacher Education*, *26*(4), 1094-1103.

Harbison, R. W., & Hanushek, E. A. (1992). *Educational Performance of the Poor: Lessons from Rural Northeast Brazil*. Oxford University Press.

Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of public economics*, *95*(7), 798-812.

Helmke, A. & Schrader, F. W. (1987). Interactional effects of instructional quality and teacher judgment accuracy on achievement. *Teaching and Teacher Education, 3*(2), 91–98.

Hickman, J. J., Fu, J., & Hill, H. C. (2012). *Technical report: Creation and dissemination of upper-elementary mathematics assessment modules.* Princeton, NJ: Educational Testing Service.

Hill, H.C., Chin, M., Blazar, D. (in preparation). *Teachers' knowledge of students: Defining a domain*.

Hill, H. C., Kapitula, L., & Umland, K. (2011). A validity argument approach to evaluating teacher value-added scores. *American Educational Research Journal*, *48*(3), 794-831.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American educational research journal*, *42*(2), 371-406.

Hoy, A. W., & Spero, R. B. (2005). Changes in teacher efficacy during the early years of teaching: A comparison of four measures. *Teaching and teacher education*, *21*(4), 343-356.

Jacob, B. A., & Lefgren, L. (2004). The impact of teacher training on student achievement quasi-experimental evidence from school reform efforts in Chicago. *Journal of Human Resources*, *39*(1), 50-79.

Justice, L. M., Mashburn, A. J., Hamre, B. K., & Pianta, R. C. (2008). Quality of language and literacy instruction in preschool classrooms serving at-risk pupils. *Early Childhood Research Quarterly*, *23*(1), 51-68.

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.

Lavy, V. (2009). Performance Pay and Teachers' Effort, Productivity, and Grading Ethics. *The American Economic Review*, 1979-2011.

Metzler, J., & Woessmann, L. (2012). The impact of teacher subject knowledge on student achievement: Evidence from within-teacher within-student variation.*Journal of Development Economics*, *99*(2), 486-496.

Mullens, J. E., Murnane, R. J., & Willett, J. B. (1996). The contribution of training and subject matter knowledge to teaching effectiveness: A multilevel analysis of longitudinal evidence from Belize. *Comparative Education Review*, 139-157.

Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of education review*,*13*(2), 125-145.

Monk, D. H., & King, J. A. (1994). Multilevel teacher resource effects in pupil performance in secondary mathematics and science. The case of teacher subject matter preparation. *Choices and consequences: Contemporary policy issues in education*, 29-58.

Muralidharan, K., & Sundararaman, V. (2011). Teacher Performance Pay: Experimental Evidence from India. *The Journal of Political Economy*, *119*(1), 39-77.

National Board for Professional Teaching Standards. (1989). *What Teachers Should Know and Be Able to Do.* Authors: Arlington, VA. Retrieved from http://www.nbpts.org/sites/default/files/what_teachers_should_know.pdf

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects?. *Educational evaluation and policy analysis*, *26*(3), 237-257.

Palardy, G. J., & Rumberger, R. W. (2008). Teacher Effectiveness in First Grade: The Importance of Background Qualifications, Attitudes, and Instructional Practices for Student Learning. *Educational Evaluation and Policy Analysis*, *30*, 111-140.

Philipp, R. A. . (2007). Mathematics teachers' beliefs and affect. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 257-315). Charlotte, NC: Information Age Publishing.

Raudenbush, S. W., Rowan, B., & Cheong, Y. F. (1992). Contextual effects on the self-perceived efficacy of high school teachers. *Sociology of Education*, 150-167.

Rice, J. K. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes*. Economic Policy Institute, 1660 L Street, NW, Suite 1200, Washington, DC 20035.

Rockoff, J. E., Jacob, B. A., Kane, T. J., & Staiger, D. O. (2011). Can you recognize an effective teacher when you recruit one?. *Education*, *6*(1), 43-74.

Rose, J. S., & Medway, F. J. (1981). Measurement of teachers' beliefs in their control over student outcome. *The journal of educational research*, *74*(3), 185-190.

Ross, J. A. (1992). Teacher efficacy and the effects of coaching on student achievement. *Canadian Journal of Education/Revue canadienne de l'education*, 51-65.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological monographs: General and applied*,*80*(1), 1.

Rowan, B., Chiang, F. S., & Miller, R. J. (1997). Using research on employees' performance to study the effects of teachers on students' achievement.*Sociology of Education*, 256-284.

Rowan, B., Correnti, R., & Miller, R. (2002). What Large-Scale Survey Research Tells Us About Teacher Effects on Student Achievement: Insights from the Prospects Study of Elementary Schools. *The Teachers College Record*, *104*(8), 1525-1567.

Rowland, T., Huckstep, P., & Thwaites, A. (2005). Elementary teachers' mathematics subject knowledge: The knowledge quartet and the case of Naomi. *Journal of Mathematics Teacher Education*, *8*(3), 255-281.

Sadler, P. M., Sonnert, G., Coyle, H. P., Cook-Smith, N., & Miller, J. L. (2013). The influence of teachers' knowledge on student learning in middle school physical science classrooms. *American Educational Research Journal*, *50*(5), 1020–1049.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational researcher*, 4-14.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard educational review*, *57*(1), 1-23.

Soodak, L.C., & Podell, D.M. (1996). Teacher efficacy: Toward the understanding of a multi-faceted construct. *Teaching and Teacher Education, 12* (4), 401-411.

Stipek, D. (2012). Context Matters. *The Elementary School Journal*, *112*(4), 590-606.

Tschannen-Moran, M., Hoy, A. W., & Hoy, W. K. (1998). Teacher Efficacy: Its Meaning and Measure. *Review of Educational Research*, *68*, 202-248.

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and teacher education*, *17*(7), 783-805.

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research*,*78*(4), 751-796.

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

Woolfolk Hoy, A., Hoy, W. K., & Davis, H. A. (2009). Teachers' self-efficacy beliefs. *Handbook of motivation at school*, 627-654.