Stanford cepa | Center for Education Policy Analysis

CEPA Working Paper No. 15-19

More than Content: The Persistent Cross-Subject Effects of English Language Arts Teachers' Instruction

AUTHORS

Benjamin Master
RAND Corporation

Susanna Loeb
Stanford University

James Wyckoff
University of Virginia

ABSTRACT

The evidence that teachers' short-term instructional effects correspond to substantial long-run impacts on students' lives provides much of the impetus for a wide range of educational reforms that are focused on identifying and responding to differences in teachers' value-added effectiveness. However, relatively little research has examined the mechanisms by which the particular knowledge or skills that teachers impart to students contribute to their longer-term success. In this paper, we investigate the persistence of teachers' value-added effects on student learning over multiple years and across subject areas. We find that the long-term instructional effects of English Language Arts (ELA) teachers are substantially more generalizable across subjects than the instructional effects of math teachers. Moreover, we find that this pattern of differential cross-subject ELA teacher effects is consistent across two different state testing regimes and district contexts. In comparison to math teachers, ELA teachers' impacts on student achievement appear to represent more broadly applicable skills that support student learning across disciplines. Our results highlight the potential for important variation in the quality of teacher-induced learning, distinct from the relative size of learning effects as measured by short-term student achievement outcomes.

VERSION

November 2015

Suggested citation: Master, B., Loeb, S., & Wyckoff, J. (2015). More than Content: The Persistent Cross-Subject Effects of English Language Arts Teachers' Instruction (CEPA Working Paper No.15-19). Retrieved from Stanford Center for Education Policy Analysis: http://cepa.stanford.edu/wp15-19

More than Content: The Persistent Cross-Subject Effects of English Language Arts Teachers' Instruction

Benjamin Master (bmaster@stanford.edu)

RAND Corporation

Susanna Loeb (sloeb@stanford.edu)

Stanford University

James Wyckoff (wyckoff@virginia.edu)

University of Virginia

This presentation/paper has not been formally reviewed and should not be distributed or cited without the author's permission.

Abstract

The evidence that teachers' short-term instructional effects correspond to substantial long-run impacts on students' lives provides much of the impetus for a wide range of educational reforms that are focused on identifying and responding to differences in teachers' value-added effectiveness. However, relatively little research has examined the mechanisms by which the particular knowledge or skills that teachers impart to students contribute to their longer-term success. In this paper, we investigate the persistence of teachers' value-added effects on student learning over multiple years and across subject areas. We find that the long-term instructional effects of English Language Arts (ELA) teachers are substantially more generalizable across subjects than the instructional effects of math teachers. Moreover, we find that this pattern of differential cross-subject ELA teacher effects is consistent across two different state testing regimes and district contexts. In comparison to math teachers, ELA teachers' impacts on student achievement appear to represent more broadly applicable skills that support student learning across disciplines. Our results highlight the potential for important variation in the quality of teacher-induced learning, distinct from the relative size of learning effects as measured by short-term student achievement outcomes.

I. Introduction

Individual K-12 teachers vary substantially in terms of their effects on students' academic performance (Rivkin, Hanushek, and Kain, 2004; Kane and Staiger, 2008; Rockoff, 2004; Nye, Konstantopoulos, and Hedges, 2004). Moreover, teachers' short-term effects on tested achievement have been shown to predict meaningful long-run outcomes in students' lives, including their college attendance and future earnings (Chetty, Friedman, and Rockoff, 2011). This link between teachers' short-term instructional effects and corresponding long-run impacts provides much of the impetus for a wide range of educational reforms focused on identifying and responding to differences in teachers' effects on students' tested achievement. However, relatively little research has examined the mechanisms by which the particular knowledge or skills that teachers impart to students contribute to their longer-term success. Accordingly, policymakers frequently focus on the relative size of teachers' short-term "value-added" impacts within a subject area, but attend less to the types of learning that teachers impart to students.

The distinction between the size and type of teachers' short term impacts may be an important one. In particular, researchers have frequently observed that reading or English language arts (ELA) teachers have smaller short-term impacts on students' relative achievement levels than math teachers (Kane and Staiger, 2008; Nye et al., 2004; Rockoff, 2004). However, Chetty and colleagues (2011) find that, while ELA teachers' effects on measured achievement levels are substantially smaller than that of math teachers, an English teacher who raises students' reading test scores by 1 unit has an impact on long-term life outcomes approximately 1.7 times that of a teacher who does the same in math. Accounting for both of these differences, the authors find that, overall, similarly ranked math and ELA teachers predict long-run effects on students' lives that are of similar magnitude.

The heterogeneous pattern of results for ELA and math teachers suggests two likely interpretations. On the one hand, teachers' measured effects on short-term student achievement might to some extent serve as a proxy for teachers' effects on other unobserved outcomes that matter to students' long-run success. Alternately, the particular academic learning reflected in teachers' value-added performance may directly influence students' long-run success, and different types of assessed learning may offer different returns over time in students' lives.

To the extent that the first interpretation is correct, a wide range of assessments may serve as interchangeable and adequate proxies for evaluating teachers' long-run instructional impacts. However, in the latter case, the particular knowledge and skills that teachers impart to students may be critically important. If some types of knowledge have more generalizable or persistent impacts on students' lives, then this can inform both our understanding of how teachers improve students' long-run outcomes, and of how best to measure teachers' contributions to student success. In particular, the available evidence suggests that the teacher-induced learning reflected in ELA achievement gains may be differentially relevant to students' future success compared to that of math achievement gains.

In this paper, we utilize data from two large urban school districts in different states to compare the generalizability and persistence of teachers' instructional effects in ELA and math. Consistent with prior research, we find that ELA and math teachers' short-term effects persist at similar rates on future assessments within the same subject area. However, we observe that the learning induced by ELA teachers has substantially more persistence across subject areas than the learning induced by math teachers. Our results suggest a mechanism by which ELA teachers may influence students' long-run outcomes to a greater degree than their short-term value-added performance would suggest. Moreover, they indicate that ELA teachers' effects on student achievement may be comparatively diffuse across subjects, and therefore may be more difficult to attribute to individual teachers than is often recognized.

II. Background

Persistence of Teacher Value-Added Effects

One way to explore the relationship between teachers' impacts on short-term and longer-term student outcomes is to examine the persistence of value-added effects on academic achievement after a student leaves a teacher's classroom. Researchers have generally observed that only a relatively small portion of teachers' proximal year effects persist and continue to impact student achievement in a subsequent school year, with most estimates ranging from between one-fifth and one-third of the initial value-added effect size (McCaffrey, Lockwood, Koretz, and Hamilton, 2004; Rothstein, 2010; Jacob, Lefgren, and Sims., 2010; Konstantopoulos and Chung, 2011; Kinsler, 2012). [1] This persistent component, however, tends to decay at a much slower rate in subsequent school years. Value-added persistence rates have generally been observed to be similar in English language arts and math.

Based on the limited duration of the majority of teachers' measured effects and the subsequent slower decay of their persistent effects, Jacob and colleagues (2010) hypothesize that persistent learning effects represent a qualitatively different type of "long-term knowledge" that teachers can impart. In this framing, long-term knowledge represents a type of "transformational learning" that has relevance to student performance in a way that is distinct from content-or test-specific short-term knowledge. The utility of long-term learning over time and across assessments also suggests that this component of teachers' effects may be particularly relevant to

[1] There is, however, some variation in estimated persistence in the literature. Using a mix of experimental and non-experimental data, Kane and Staiger (2008) estimate persistence of value added effects that are close to 50% after one year, while Lockwood, et al. (2007) estimate persistence parameters that are less than 1/5th.

students' long-run life outcomes. Researchers have generally observed only modest correlations between teachers' short-and-long-term value-added effects (Rothstein, 2010; Mariano, McCaffrey, and Lockwood, 2010)[2], which indicates that there may be meaningful heterogeneity in the quality or types of knowledge that individual teachers impart to students.

No extant research that we are aware of has examined the persistence of teacher-induced learning effects across subjects. However, the persistence of long-term knowledge across subjects may offer valuable insight into the mechanisms by which teachers influence students' longer term outcomes. To the extent that learning in different subject areas persists at comparable rates across-subjects, this would support the notion that teachers' instructional effects, on average, include a comparable portion of generalizable knowledge that is useful in any academic context. Alternately, if learning of certain types is differentially persistent across subjects, then this would imply that the subject matter of teachers' instruction corresponds to variation in its generalizability and future impact. This could be the case, for instance, if teacher-induced improvements in students' reading skills promote their ability to acquire knowledge in a wider array of subjects.

Cross-subject Instructional Effects

A growing body of evidence indicates that teachers can have meaningful effects on their students' contemporaneous performance across subject areas, although much of this evidence comes from studies at the high school level. For example, Aaronson, Barrow, and Sander (2007) observe that 9th grade math and ELA teachers in Chicago Public Schools have sizeable cross-subject effects in addition to their within-subject effects. Buddin and Zamarro (2009) observe the

[2] Across different value-added models and datasets, researchers have identified correlations ranging from 0.3 to 0.6 between teachers' proximal and future-year effects.

same trend in students in grades 9-11 in the Los Angeles Unified School District. Koedel (2009) finds that 9-11th grade teachers in both ELA and math in the San Diego Unified School District impact student performance on students' reading achievement, although science and social studies teachers do not. Jackson (2012), however, does not find evidence of significant cross-subject effects of 9th grade Algebra and English teachers in North Carolina.

Both theory and some empirical evidence indicate that ELA instructional effects may be especially generalizable across subjects. For example, students' reading and language skills have been shown to be important across a range of other subjects and may be particularly important for students of lower socio-economic status or those with limited proficiency in English (Abedi and Lord, 2001; O'Reilly and McNamara, 2007; Chang, Singh, and Filer, 2009). Moreover, in a recent study of teachers' cross-subject effects Yuan (2014) examines teachers in middle school grades in an anonymous urban school district and finds that ELA teachers contribute to student achievement in mathematics, ELA, science, and social studies, while, by comparison, mathematics teachers contribute significant cross-subject effects only in ELA achievement.

Teachers' contemporaneous cross-subject effects may stem from a variety of mechanisms. For example, teachers may directly collaborate in preparation or instruction, improving peer performance or reinforcing a common set of skills. Similarly, the content of some curricula may explicitly overlap to varying degrees and reflect common knowledge that is taught across subjects and classrooms. Alternately, knowledge generated in one subject area may consist of generalizable skills that directly support students' learning in an otherwise unrelated subject area. Thus, it is difficult to disentangle the different ways in which teachers may be contributing to their peers' contemporaneous cross-subject performance.

Contribution

In this study, we expand upon the prior research related to both teachers' persistent effects and to teachers' cross-subject effects. Following a conceptual and methodological framework described by Jacob and colleagues (2010), we investigate the persistence of teachers' value-added effects, distinguishing between short-term, test-specific knowledge and longer-term learning that accumulates. We extend this framework by differentiating between long-term knowledge that is content-specific and generic knowledge that persists across subjects.

Our approach allows us to identify the persistent impacts of the subject-specific learning that prior-year teachers impart to students. As a consequence, we are better able to isolate the extent to which different types of teacher-induced learning generalize across subject areas, distinct from contemporaneous teacher spillover effects, which may stem from multiple factors. Finally, in order to gauge the generalizability of our findings, we consider evidence from two large urban school districts that are located in different states and that utilize different standardized assessments of students' performance in English language arts and math.

We specifically consider the following research questions:

1) What is the rate of persistence of previously assessed ELA and math knowledge within and across subjects?
2) What is the rate of persistence of teachers' value-added effects on student learning, within and across subjects?
3) How does the relative magnitude of ELA and math teachers' effects on achievement differ when considering short-term, subject-specific effects versus long-term, multi-subject effects?

The remainder of the paper proceeds as follows. In the following section III, we describe the data utilized in the study. Section IV details our methods for estimating value-added measures and value-added persistence within and across subjects. In section V we detail our results, and we conclude in section VI with a discussion of conclusions and potential limitations.

III. Data

Administrative Data

In order to investigate the within and across-subject persistence of teachers' value-added effects, we draw upon extensive administrative data about students, teachers, classrooms, and schools in two large urban school districts: New York City (NYC) and Miami-Dade County Public Schools (MDCPS). In both NYC and MDCPS, our available data includes students in third through eighth grade from school years (SY) 2003-04 through SY 2011-12. Both district data sets include data on students' annual achievement test scores in ELA and math, and identification of students' primary teacher and classroom in each year and subject area. For the purposes of our analysis, we standardize students' achievement test scores within each grade, subject, and year in each sample.

In addition to achievement data, we also have access to a rich set of demographic and behavioral characteristics that we utilize in our analyses. For both NYC and MDCPS students, these characteristics include their race/ethnicity and home language, as well as their absences, suspensions, school transfers, free or reduced price lunch status, disability or special education status, and English language learner status in each school year.

We estimate teacher value-added effects using a sample of student-year records in grades four through eight for whom current and prior year achievement data is available. However, because we are investigating the persistence of teachers' value-added effects on student achievement in periods *after* they teach a student, the sub-samples for our persistence analyses are restricted to student-year observations in which we can identify prior-year teachers with appropriate value-added scores, as detailed in the Methods section below. In practice, these requirements reduce our analysis of teachers' persistent effects on student outcomes to student-year observations in grades five through eight, from SY 2005-06 through SY 2011-12. This sample includes only students who are present in each district for at least three consecutive years.[3] Summary statistics for each district's analytical sample are presented in Table 1.

TABLE 1

Summary statistics for students, teachers, and schools in the analytical sample for each district

Variables	New York City	Miami-Dade
A. Students		
% Free or Reduced price lunch	71.8	66.0
% Black	30.2	25.0
% Hispanic	36.3	63.0
% White	16.4	9.5
% Female	51.8	50.2
N of Distinct Students	473,004	180,512
B. Teachers		
N of Distinct ELA Teachers	13,660	5,016
N of Distinct Math Teachers	13,368	4,363
C. Schools		
Average % of students eligible for free lunch	69.1 (24.0)	66.3 (22.1)
Average % of students Black	36.5 (30.1)	25.1 (30.3)
Average % of students Hispanic	39.3 (26.2)	63.0 (29.22)
N of Distinct Schools	1,169	359

[3] Note that when evaluating teachers' 2-year persistent effects on student outcomes, we utilize a more restricted sample consisting of only observations of students who are present in the district for at least four consecutive years, and who are in grades six through eight, from SY 2006-07 through SY 2011-12.

Note: Analytical samples consists of students in grades 5 through 8 in school years 2005-06 through 2010-11.

IV. Methods

Teacher Value-Added Measures

In order to examine the persistence of teachers' effects, we first generate teacher value-added estimates of the effects of each teacher on tested student achievement in each year. We employ a value-added model that has been used by the NYC Department of Education in the past to evaluate their teachers' performance (Value-Added Research Center, 2010). Conceptually, this model compares teachers to other "similarly circumstanced" teachers by first predicting students' achievement with both prior achievement measures and a range of observable student, classroom, and school characteristics that may influence their achievement, and then attributing the remaining unexplained variation in student performance to individual teachers. Details of our value-added model specification are provided in Appendix A.

Estimating the Persistence of Teacher Value-Added Effects

We estimate the persistence of teachers' value-added effects using an instrumental variables approach described by Jacob et. al (2010). As previously discussed, these authors conceptualize students' tested knowledge as a combination of "short term" knowledge that has no observed impact on future achievement, and "long-term" knowledge that is relevant to both contemporaneous and future achievement tests. In their formulation, observed student achievement Y in a given period t represents a combination of that student's long-term knowledge from a prior period and all contemporaneous impacts (including teachers' effects) that influence both their long and short term knowledge in the current period:

$$Y_t = \theta \mathrm{y}_{l,t-1} + \eta_t^l + \eta_t^s \qquad (1)$$

Here, current achievement is a function of contemporaneous impacts η_t^l and η_t^s on long and short term knowledge, as well as long-term knowledge in the prior period $y_{l,t-1}$, which carries forward with some rate of decay (1-θ).

In practice, we do not directly observe long-term knowledge, but rather the sum of long-term and short-term knowledge assessed in the prior period, Y_{t-1}. In light of this, the authors describe how an ordinary least squares (OLS) coefficient θ_{OLS} for a regression of current achievement on prior achievement converges to the following:

$$\text{plim}\left(\hat{\theta}_{OLS}\right) = \theta\left(\frac{\sigma_{yl}^2}{\sigma_{yl}^2+\sigma_{ys}^2}\right) \qquad (2)$$

This equation shows that because prior knowledge consists of a mix of long- and short-term knowledge, the OLS coefficient will be attenuated to the extent that Y_{t-1} consists of short-term, rather than long-term knowledge. In lieu of an OLS estimate of the persistence of observed knowledge, Jacob and colleagues (2010) use an instrumental variables approach to estimate the decay of prior long-term knowledge, using twice lagged achievement Y_{t-2} as an instrument for Y_{t-1}. This estimator, which we refer to as $\hat{\theta}_{LT}$, purges Y_{t-1} of its short-term knowledge component. In practice the vast majority of a student's previously assessed long-term knowledge persists between one year and the next, with a value of $\hat{\theta}_{LT}$ close to 1. This serves as a benchmark for our subsequent estimation of teachers' contributions to long-term knowledge.

Following a similar approach, we estimate the proportion of a teacher's effect that consists of long-term knowledge by instrumenting each student's lagged knowledge Y_{t-1} with their lagged teacher's contribution (value-added) to that knowledge. The lagged teacher's total contribution to a student's lagged knowledge is a combination of his or her contribution to long- and short-term lagged knowledge, expressed as $\mathrm{M}_{t-1} = \mu^{l}_{t-1} + \mu^{s}_{t-1}$. Thus, the second stage estimator $\hat{\theta}_M$ converges to:

$$\text{plim}(\hat{\theta}_M) = \theta\left(\frac{\sigma^2_{\mu l}}{\sigma^2_{\mu l}+\sigma^2_{\mu s}}\right) \tag{3}$$

Given an estimate of θ that is close to 1, $\hat{\theta}_M$ approximates the fraction of teacher value-added that is attributable to long-term, rather than short-term, knowledge creation.

In practice, student assignment to teachers is nonrandom, and therefore the measured quality of a student's lagged teacher may be correlated with the quality of their current teacher. To minimize possible bias in our teacher persistence estimates due to persistent nonrandom patterns of teacher assignment, we include in our instrumental regression to estimate $\hat{\theta}_M$ additional controls for both student level covariates χ and for contemporaneous classroom fixed effects π (which incorporate school, year and grade fixed effects). In addition, because teachers' value-added scores in any given year include estimation error that is correlated with other classroom-specific learning shocks in that year, we calculate, for each student in each period, their lagged teachers' average value-added score across all years other than the year in which they taught that student, expressed as $\mathrm{T}_{ijt-1} = \sum_{y\neq t-1} \mathrm{M}_{jy}$. The second-stage equation for estimating the persistence of teacher value added then becomes:

$$Y_{icjt} = \theta_{\mathrm{M}} Y_{it-1} + X_{it} + \pi_{cjt} + \varepsilon_{ijt} \tag{4}$$

Where the values of T_{ijt-1} for the lagged teachers serve as the excluded instruments for students' prior test scores in the first stage. In this formulation, persistence is a function of variation in the measured quality of the lagged teacher, distinct from the effects of the student's teacher or school in the current year.[4]

[4] Jacob et al. (2010) note, however, that our estimates of persistence may still be biased if schools adjust the instructional inputs (other than classroom assignments) that students receive, as a response to the quality of their

Finally, in order to estimate teachers' cross-subject persistence, we modify equation 4 by replacing our outcome measure, Y_{icjt}, with a student's achievement in the alternate subject, Y_{icjt}^{alt}. Thus, for example, we model students' current math achievement as a function of their prior-year ELA achievement, instrumented by their lagged ELA teacher's measured value-added ability. In addition, when predicting current math achievement, we include classroom fixed effects corresponding to their current-year math classroom assignment, rather than their ELA classroom assignment. We do the reverse when estimating the persistence of lagged math teachers' value-added on students' current ELA achievement. This allows us to estimate the proportion of teachers' value added effects in one subject that persist and influence cross-subject outcomes in a future school year, while controlling for contemporaneous contributions to students' achievement that might be correlated with lagged teacher quality. We make the same adjustments when estimating the general rate of persistence of previously assessed long-term knowledge across subjects.

V. Results

Persistence of Knowledge within and Across Subjects

We begin by investigating the persistence of long-term knowledge within subjects, via a two-stage regression in which prior achievement is instrumented with twice-lagged achievement, as previously discussed. These results are shown in Table 2. Consistent with prior research and with our intuitive understanding of long-term knowledge, we find that nearly all of previously assessed long-term knowledge (i.e. knowledge that is relevant across two prior school years) also persists into a third year. Across NYC and MDCPS, partial coefficients on long-term knowledge range in a fairly narrow band from 0.931 to 1.057. These results are in contrast to the coefficients

---

lagged teacher. This could occur, for instance, if effective teacher raise students' achievement and this in turn leads schools to provide fewer instructional supports to the student.

on students' observed prior-year test scores, which reflect a mix of short-and-long-term knowledge (including measurement error) and persist at substantially lower rates.

TABLE 2

Estimates for the Persistence of Observed Knowledge and Long Term Knowledge, Within Subject

	NYC		MDCPS	
	Observed Knowledge	Long Term Knowledge	Observed Knowledge	Long Term Knowledge
Predicting Math with Math Knowledge				
Coefficient on lagged achievement	0.771	0.931	0.794	1.015
	(0.001)	(0.002)	(0.001)	(0.006)
First-stage F-statistic	-	169317	-	31454
N of student-year observations	983564		337195	
Predicting ELA with ELA Knowledge				
Coefficient on lagged achievement	0.645	0.946	0.775	1.057
	(0.001)	(0.003)	(0.001)	(0.006)
First-stage F-statistic	-	75008	-	31916
N of student-year observations	936062		342564	

Note: Coefficient for Observed Knowledge from a regression of current achievement on prior achievement. Coefficient for Long Term Knowledge from an instrumental variables (IV) regression of current achievement on prior achievement instrumented with twice-lagged achievement, with additional controls for student characteristics and contemporaneous classroom fixed effects.

Next, in Table 3, we show how previously assessed long-term knowledge in one subject area predicts current performance in the alternate subject. In comparison to our within-subject results, we find that knowledge that is relevant over time and across assessments in one subject area persists at a substantial, but reduced rate into the alternate subject. Coefficients when predicting math with ELA knowledge or ELA with math knowledge range from 0.614 to 0.713. Across both NYC and MDCPS, ELA long-term knowledge persists at a somewhat higher rate (0.639 and 0.713, respectively) across subjects than math long-term knowledge (0.614 and 0.687). These results suggest that much of the overall long-term knowledge that students possess reflects foundational skills, abilities, or other human capital that are relevant across academic subjects. However, a meaningful portion appears to consist of subject-specific skills as well.

TABLE 3

Estimates for the Persistence of Observed Knowledge and Long Term Knowledge, Across Subjects

	NYC		MDCPS	
	Observed Knowledge	Long Term Knowledge	Observed Knowledge	Long Term Knowledge
Predicting Math with ELA Knowledge				
Coefficient on lagged achievement	0.581	0.639	0.664	0.713
	(0.001)	(0.003)	(0.001)	(0.006)
First-stage F-statistic	-	73089	-	31916
N of student-year observations	932966		342564	
Predicting ELA with Math Knowledge				
Coefficient on lagged achievement	0.609	0.614	0.664	0.687
	(0.001)	(0.002)	(0.001)	(0.006)
First-stage F-statistic	-	170557	-	31454
N of student-year observations	976801		337195	

Note: Coefficient for Observed Knowledge from a regression of current achievement on prior achievement. Coefficient for Long Term Knowledge from an instrumental variables (IV) regression of current achievement on prior achievement instrumented with twice-lagged achievement, with additional controls for student characteristics and contemporaneous classroom fixed effects. .

The Persistence of Teacher-induced Learning within and Across Subjects

To address our second research question, we first examine the extent to which teachers' value-added effects on student learning persist over time (i.e. consist of long-term knowledge) within each subject area. In table 4, we show results from a 2-stage regression in which we predict current achievement in each subject area with students' prior achievement in the same subject, instrumented by their prior same-subject-teacher's estimated value added. In NYC, we find that only around 20 percent of teachers' value-added effects persist into a subsequent school year, with similar results across math and ELA teachers. 1-year persistence rates are higher in

MDCPS in both subjects, but are also fairly similar across math (0.408) and ELA (0.362). Estimates from MDCPS are based on a substantially smaller sample of students than those in NYC, and standard errors are generally larger in both subjects.

Differences in the magnitude of value-added persistence across districts may stem from a range of factors, including differences in the year-over-year alignment and content covered by state tests in each district. Overall, however, our estimates of same-subject teacher persistence reflect a comparable range to those reported in prior research, and align with previous research showing roughly equivalent persistence rates of teacher-induced learning within each subject area.

TABLE 4

Estimates for the Persistence of Teachers' Value-Added Effects, Within-Subject

	NYC		MDCPS	
	Teacher Value Added Effects, 1-year	Teacher Value Added Effects, 2-year	Teacher Value Added Effects, 1-year	Teacher Value Added Effects, 2-year
Predicting Math with Prior Effects on Math				
Coefficient on lagged achievement	0.195	0.078	0.362	0.195
	(0.008)	(0.009)	(0.016)	(0.017)
First-stage F-statistic	8725	9187	2237	2493
N of student-year observations	981161	503125	335448	155857
Predicting ELA with Prior Effects on ELA				
Coefficient on lagged achievement	0.194	0.169	0.408	0.334
	(0.018)	(0.008)	(0.038)	(0.040)
First-stage F-statistic	2111	9764	457	482
N of student-year observations	933185	524711	340379	162450

Note: Coefficients for Teacher Value Added Effects from regressions of current achievement on prior achievement instrumented with prior year (or twice-lagged) teacher value added quality. Models include controls for current student characteristics and classroom fixed effects from the test subject of the dependent variable. 1-year value added effects reflect initial teacher effects in grades 4-7 and outcomes in grade 5-8. 2-year value added effects reflect initial teacher effects in grades 4-6 and persistence into grades 6-8.

Results for 2-year persistence in the same subject vary somewhat across subjects and districts, but are also generally consistent with results from prior research. While the majority of teachers' value-added effects do not persist 1 year after instruction, persistence estimates 2 years after instruction suggest a somewhat slower rate of decay in teachers' effects. It is important to note, however, that our estimates of teachers' 2-year value added persistence are based on a different sample of teachers, students, and grade-levels, and thus cannot be directly compared to the 1-year teacher persistence estimates.

In Table 5, we show the cross-subject persistence of teachers' value-added effects on student learning in each district. In other words, we predict current achievement in math or ELA as a function of prior achievement in the alternate subject, instrumented by the student's prior-teacher in the alternate subject. Across both NYC and MDCPS, we observe stark differences in the rate of cross-subject persistence between ELA and math teachers, with ELA teacher effects persisting at a much higher rate across subjects. For example, in NYC, ELA teachers' persistence coefficient across subject is 0.163, which is more than 80 percent of their persistence rate within the same subject (0.195). In contrast, NYC math teachers' persistence coefficient across subjects is 0.044, less than a quarter of their same-subject persistence (0.194).

The differential cross-subject persistence of ELA teachers' instruction is also apparent in MDCPS. In this district, ELA teachers' estimated cross-subject persistence rate (0.418) is virtually identical to their within-subject persistence (0.408). In contrast, MDCPS math teachers' cross subject persistence rate (0.104) is less than 30 percent of their estimated within-subject persistence (0.362).

We observe similar, albeit less dramatic, trends of differential cross-subject persistence of teacher-induced learning in ELA 2 years after instruction. For example, in NYC the estimated 2-year within-subject ELA persistence rate is 0.169, which is roughly twice that of their persistence across subjects (0.081). In math, 2-year cross-subject persistence (0.021) is just over a quarter of the within-subject persistence estimate (0.078). In MDCPS, ELA teachers' two-year cross-subject persistence (0.240) is roughly two thirds of their within-subject persistence (0.362), while math teachers' two-year cross subject persistence (0.090) is less than half of their 2-year persistence estimates within the same subject (0.195).

TABLE 5

Estimates of the Persistence of Teachers' Value-Added Effects Across Subjects

	NYC		MDCPS	
	Teacher Value Added Effects, 1-year	Teacher Value Added Effects, 2-year	Teacher Value Added Effects, 1-year	Teacher Value Added Effects, 2-year
Predicting Math with Prior Effects on ELA				
Coefficient on lagged achievement	0.163	0.081	0.418	0.240
	(0.017)	(0.008)	(0.038)	(0.041)
First-stage F-statistic	2104	9764	457	410
N of student-year observations	914540	513606	317485	146359
Predicting ELA with Prior Effects on Math				
Coefficient on lagged achievement	0.044	0.021	0.104	0.090
	(0.010)	(0.010)	(0.017)	(0.017)
First-stage F-statistic	8626	9049	2823	2586
N of student-year	959327	493501	318537	145791

observations

Note: Coefficients for Teacher Value Added Effects from regressions of current achievement on prior achievement instrumented with prior year (or twice-lagged) teacher value added quality. Models include controls for current student characteristics and classroom fixed effects from the test subject of the dependent variable. 1-year value added effects reflect initial teacher effects in grades 4-7 and outcomes in grade 5-8. 2-year value added effects reflect initial teacher effects in grades 4-6 and persistence into grades 6-8.

Overall, our analysis of the persistence of teacher-induced learning indicates that ELA teachers' impacts on students' assessed ELA skills are substantially more persistent in math than math teachers' impacts on future-year ELA performance. The differential relevance of learning in ELA suggests that ELA teachers' instruction touches upon more generalizable skills than that of math teachers.

Comparing the Relative Magnitude of Teacher-induced Learning in ELA and Math

While the previous results indicate that ELA and math teachers' effects persist at very different rates across subjects, this does not necessarily reflect the relative magnitude of their long-term achievement impacts. For instance, while ELA teachers' effects appear to be more generalizable across subjects, they are also frequently smaller in magnitude. In table 6, we provide the first-stage partial coefficients from our estimates of teachers' 1-year and 2-year within-subject persistence. We use these to estimate the size of teachers' initial and persistent instructional impacts on student learning across each district and each analytical sample. In this table, teachers' measured value added is standardized to have a standard deviation of 1 and a mean of 0.

For the initial year of instruction, we are able to measure teachers' impact only within a single subject area. Using this criteria to measure instructional impact, we observe that, across three of our four samples, teachers' value-added ability predicts substantially larger instructional impacts in math than in ELA. Math and ELA teacher effect sizes are roughly equivalent in our 2-year persistence sample in NYC, however. In other words, when we consider only proximal year effects in their primary subject, the measured instructional impact of being assigned a high-quality ELA teacher generally appears much smaller than that of a similarly ranked math teacher.

When considering teachers' persistent effects and accounting for both within-and-across subject impacts, however, the relative magnitude of ELA teachers' instruction grows substantially. While math teachers' persistent effects on both math and ELA continue to be somewhat larger than ELA teachers' dual-subject effects in most cases, including cross-subject effects in an additional subject area goes a long way towards equalizing the impact of teachers in each subject area. For instance, in our 1-year persistence sample in NYC, ELA teachers' proximal within subject effects are estimated to be roughly half (52.4%) of that of math teachers. However, within the same sample, ELA teachers' persistent, multi-subject effects are much closer in size to that of math teachers (78.4%). In other words, while being assigned to a higher quality ELA teacher may appear to be substantially less impactful than assignment to a higher quality math teacher in terms of within-subject effects, assignment to a higher-quality ELA teacher is much more impactful when we consider their long-term, cross-subject impacts, rather than their short-term, within-subject effects.

TABLE 6
Relative magnitude of math and ELA teachers' estimated achievement impacts within a single subject in the year of instruction and in future periods across two subjects

	NYC	MDCPS
First Stage Coefficients, 1-Year Persistence Sample		
Math teacher VAM on math	0.109	0.079
ELA teacher VAM on ELA	0.057	0.033
Size of within-subject proximal ELA effects relative to math	52.4%	42.1%
Estimated 1-Year Persistance		
Math VAM on math	0.021	0.029
Math VAM on ELA	0.005	0.008
ELA VAM on ELA	0.011	0.014
ELA VAM on math	0.009	0.014
Size of dual-subject persistent ELA effects relative to math	78.5%	74.6%
First Stage Coefficients, 2-Year Persistence Sample		
Math teacher VAM on math	0.123	0.095
ELA teacher VAM on ELA	0.135	0.040
Size of within-subject proximal ELA effects relative to math	109.8%	42.1%
Estimated 2-Year Persistance		
Math VAM on math	0.010	0.019
Math VAM on ELA	0.003	0.009
ELA VAM on ELA	0.023	0.013
ELA VAM on math	0.011	0.010
Size of dual-subject persistent ELA effects relative to math	277.2%	84.8%

Note: Value-added measures computed as the mean across teachers' value added in all years other than the year they taught the student, and standardized across a student-level sample to have a mean of 0 and a standard deviation of 1. First stage coefficients are from a two-stage regression that includes controls for current student characteristics and classroom fixed effects from the test subject of the dependent variable. Implied persistence is a function of first stage coefficients and the estimated 1-year and 2-year persistence rates of teachers' value added effects as reported in Tables 4 and 5.

VI. Conclusion and Discussion

This study offers new insight into k-12 teachers' instructional effects, contributing to our understanding of how teachers support student outcomes over time. We identify important differences in the type of learning that teachers of different subjects impart, as evidenced by a much higher rate of cross-subject persistence associated with teacher-induced learning in ELA than in math. These results may help to explain the mechanisms by which effective ELA teachers contribute similarly to students' long-run outcomes as effective math teachers, in spite of their smaller contemporaneous effects on within-subject achievement (Chetty et al., 2011). While teachers' contemporaneous effects on student ELA achievement are frequently smaller than math teachers' effects, the knowledge that they impart appears to be differentially important to student learning more broadly, and the size of their effects, when measured in terms of long-term knowledge over multiple subjects, are much more comparable with that of math teachers. Overall, our findings corroborate prior theory and research regarding the importance of students' reading skills across a variety of contexts (Abedi and Lord, 2001; O'Reilly and McNamara, 2007; Chang et al., 2009).

The results of this study also add to a growing body of evidence regarding teachers' cross-subject spillover effects. While previous studies provide evidence of substantial contemporaneous spillover effects across different teachers' classrooms (Yuan, 2014; Aaronson et al., 2007; Buddin and Zamarro, 2009; Koedel, 2009), these effects may reflect a variety of contributing factors, including explicit peer-to-peer collaboration or overlapping content coverage. In contrast, our methods allow us to better isolate the persistent spillover effects of different types of student knowledge, and we demonstrate that ELA learning in particular is more broadly relevant to performance in another subject area. Our study is also among the first to explore the cross-subject effects of teachers in elementary and secondary school grade levels. Consistent with Yuan's (2014) findings for teachers in secondary grades, we find greater spillover from ELA teachers' instruction, relative to that of math teachers.

To the extent that teachers' instructional effects influence student achievement across subject areas, districts may miss valuable information if they rely only on within-subject student learning to evaluate teacher performance. We find that ELA teachers' contributions to student learning are particularly diffuse, and thus their contributions may be captured less accurately by typical within-subject value-added measures of teacher performance. Our results highlight the potential importance of attending to team-level indicators of teachers' contributions to student achievement, or to estimating value-added models that simultaneously control for multiple different subject-teachers' contributions.

A key limitation of our analysis, and of most data sets that have been used to estimate teacher value-added performance, is that we are unable to examine the import of instructional effects in subject areas other than ELA and math. As a result, we are unable to examine whether the cross-subject generalizability of ELA learning extends to additional subject areas other than math. While we hypothesize that this might be the case, future research is needed to test this.

Another important caveat when measuring the persistence of teacher induced learning is that persistence rates may vary as a function of the overlap in assessed content between one year's test administration and the next. Similarly, cross-subject effects may differ in magnitude as a function of the overlap in content between different subject area tests, particularly if some non-ELA tests require student reading and writing skills to a greater degree than others. That said, the consistency of our main findings across two large urban school districts that utilize different tests and content standards provides valuable evidence that our findings are likely to generalize across a wide range of K-12 school settings.

K-12 teachers have substantial impacts on students' academic achievement, and our best evidence to-date indicates that these impacts are predictive of meaningful long-run effects on students' life outcomes (Chetty et al., 2011). This predictive utility of value-added measures reinforces the use of short-term student achievement outcomes as valid indicators of teachers' contributions. However, it is important for researchers to also understand the mechanisms by which short term instructional effects translate into longer-term student outcomes. Without a clearer understanding of these mechanisms, we may be unable to differentiate between short-term effects that are transient or context-dependent, versus those that have broader benefits for students. The findings from this study highlight the potential for meaningful variation in the type of learning that different teachers impart, as evidenced by substantial differences in how different instructional effects play out over time. Additional research is needed to further illuminate the processes by which different types of learning yield persistent benefits for students.

Appendix A –Value-Added Model

Following University of Wisconsin (2010), we compute teacher-by-year value-added scores in three stages. In the first stage, we estimate the coefficients λ for students' pretests and β for student-level characteristics on students' posttest scores. To estimate these coefficients, we regress posttest Y_t of student i in classroom c with teacher j in school s at time t on their same-subject pretest Y_{t-1}, other-subject pretest Y_{t-1}^{alt}, a vector of student-level time varying and time invariant variables X, and a set of indicator variables representing individual classroom fixed effects π, which can be expressed as:

$$Y_{icjst} = \lambda Y_{it-1} + \lambda^{alt} Y_{it-1}^{alt} + \beta X_{it} + \pi_{cjst} + \varepsilon_{icjst} \qquad (5)$$

Our student-level characteristics include students' gender, race, an indicator for whether the student's home language is English, student eligibility for free or for reduced price lunch, student disability status, English language learner status, an indicator for whether the student switched schools in the prior year, and the number of prior-year absences for the student. Because the effects of characteristics may vary across grade levels, we also include interactions of each student characteristic with each individual grade level.

The first-stage regression is estimated using an errors-in-variables approach (following Fuller (*22*)) that accounts for measurement error in pretests Y_{it-1} and Y_{it-1}^{alt}. This removes the variance in the pretests that is attributable to measurement error. To facilitate this approach, we rely on reliability information as reported in the technical manuals for the state assessments in New York and Florida.

In the second stage, we use the estimated coefficients λ and β from our first stage to compute a new left-hand side variable q_{icjst}, where $q_{icjst} = Y_{icjst} - \lambda Y_{it-1} - \lambda^{alt} Y_{it-1}^{alt} - \beta X_{icjst}$. q_{icjst} is, then, the difference between the student's actual score and what we would predict it to be given background characteristics and prior performance. We then regress q_{icjst} on a vector C of classroom-level characteristics, time-varying school-level characteristics K, and individual year and grade dummy indicators:

$$q_{icjst} = \gamma C_{cjst} + \eta K_{st} + \alpha_t + \rho_g + w_{icjst} \qquad (6)$$

Classroom-level characteristics include the racial and home language composition of the classroom, class size, the percent of students who are free or reduced price lunch eligible, percent of students who are English language learners, the class average number of prior year absences, the class average prior year test scores in the same and alternate subject, and the standard deviation of classroom test scores in each subject. As we did for the student covariates, we include interactions of each classroom characteristic with each grade level indicator. School characteristics include total enrollment, the percent of black, white, and Hispanic students in the school, and a control for the percent of students eligible for free or reduced price lunch. When running this regression, we specify a classroom random effect to take into account that errors are correlated within classrooms. From this regression, we obtain an estimate of w_{icjst}, that represents the residual test score variation for each student in each year that is not explained by our observable student, classroom, or school characteristics.

In our third stage, we estimate individual teacher value-added measures in each year, τ_{jt}, by attributing all remaining variation in students' post-test scores to a combination of the individual teacher effects and error. This can be expressed as

$$w_{icjst} = \tau_{jt} + \varepsilon_{icjst} \qquad (7)$$

We obtain estimates of the error term ε_{icjst} by subtracting each teachers' mean effect, τ_{jt}, from the estimates of w_{icjst}. Finally, we standardize our teacher-by-year effect estimates across our sample to have a mean of zero and a standard deviation of one[5]. We include in our analysis only teacher-by-year effects that are based on at least 5 students and fewer than 100 students.

[5] Prior to standardizing, the standard deviation of our (un-shrunken) ELA teacher-by-year value added measures is 0.23, while the standard deviation for math teachers' value added measures is 0.27.

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.

Abedi, J. & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education, 14*(3), 219–234. doi:10.1207/S15324818AME1403_2

Buddin, R. & Zamarro, G. (2009). *Teacher effectiveness in urban high schools* (WR-693-IES). Santa Monica,CA: RAND Corporation. Retrieved from http://www.rand.org/content/dam/rand/pubs/working_papers/2009/RAND_WR693.pdf

Chang, M., Singh, K., & Filer, K. (2009). Language factors associated with achievement grouping in math classrooms: A cross-sectional and longitudinal study. *School Effectiveness and School Improvement, 20*(1), 27–45. doi:10.1080/09243450802605704

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2011). *The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood* (No. w17699). National Bureau of Economic Research.

Jackson, C. K. (2012). *Non-cognitive ability, test scores, and teacher quality: Evidence from 9th grade teachers in North Carolina* (Working Paper 18624). Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w18624

Jacob, Brian A., Lars Lefgren, and David P. Sims. 2010. The persistence of teacher-induced Learning. *Journal of Human Resources, 45*(4): 915-943.

Kane, T. J., & Staiger, D. O. (2008). *Estimating teacher impacts on student achievement: An experimental evaluation* (No. w14607). National Bureau of Economic Research.

Kinsler, J. (2012). Beyond Levels and Growth Estimating Teacher Value-Added and its Persistence. *Journal of Human Resources*, *47*(3), 722-753.

Koedel, C. (2009). An empirical analysis of teacher spillover effects in secondary school. *Economics of Education Review*, *28*, 682–692. doi:10.1016/j.econedurev.2009.02.003

Konstantopoulos, S., & Chung, V. (2011). The persistence of teacher effects in elementary grades. *American Educational Research Journal*, *48*(2), 361-386.

This presentation/paper has not been formally reviewed and should not be distributed or cited without the author's permission.

Lockwood, J. R., McCaffrey, D. F., Mariano, L. T., & Setodji, C. (2007). Bayesian methods for scalable multivariate value-added assessment. *Journal of Educational and Behavioral Statistics*, *32*(2), 125-150.

Mariano, L.T, McCaffrey, D.F., Lockwood, J.R. (2010). A model for teacher effects from longitudinal data without assuming vertical scaling. *Journal of educational and behavioral statistics*, *35*(3), 253-279.

McCaffrey, D. F., Lockwood, J. R., Koretz, D., Louis, T. A., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of educational and behavioral statistics*, *29*(1), 67-101.

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

O'Reilly, T., & McNamara, D. (2007). The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional "high-stakes" measures of high school students' science achievement. *American Educational Research Journal, 44*(1), 161–196. doi:10.3102/0002831206298171

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*(2), 417-458.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review*, *94*(2), 247-252.

Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. *The Quarterly Journal of Economics*, *125*(1), 175-214.

Value-Added Research Center. (2010) "NYC Teacher Data Initiative: Technical Report on the NYC Value-Added Model." Madison Wisconsin: Wisconsin Center for Education Research. Retrieved February 13, 2015 from: http://schools.nyc.gov/NR/rdonlyres/A62750A4-B5F5-43C7-B9A3-F2B55CDF8949/87046/TDINYCTechnicalReportFinal072010.pdf

Fuller, W. A. (2009). *Measurement error models* (Vol. 305). John Wiley & Sons.