

Improvements in Math Instruction and Student Achievement Through Professional Learning Around the Common Core State Standards in Chicago

Elaine Allensworth 

Sarah Cashdollar 

Julia Gwynne

University of Chicago

Existing literature on the impact of Common Core State Standards in Math has shown little benefit, but it has not examined variation in outcomes based on implementation strategies, student subgroups, or outcomes other than test scores. We use a difference-in-differences approach with school fixed effects to compare outcomes in pre- and poststandards years across schools with different levels of participation in professional learning around the standards in the middle grades in Chicago. Postimplementation, there were significantly greater improvements in student reports of standards-aligned instructional practices, math grades, pass rates, and test scores in schools with more extensive professional learning around the standards, among students with low and average initial achievement. Relationships were largely not significant for students with high initial achievement. We discuss why Chicago might have seen positive results, including the district emphasis on professional learning around the practice standards and differential impacts based on student prior achievement.

Keywords: *educational policy, instructional practices, longitudinal studies, mathematics education, middle schools, policy analysis, professional development, quasi-experimental analysis*

THE adoption of the Common Core State Standards in Mathematics (CCSS-M) by most states across the country brought ambitious new targets for what students should learn and be able to demonstrate in school, and significant new demands on teachers. Almost 10 years after states began implementing the CCSS-M, there is only limited research on how student outcomes have changed in response to the new standards. The studies that exist focus exclusively on test scores and do not differentiate effects for students with different achievement levels. They also tell us little about which strategies for implementing the standards might be effective, as they examine average changes across entire states.

New standards by themselves cannot raise student achievement—they depend on the work of districts, schools, and teachers to interpret the standards and effectively shift instructional practice (Toch, 1991; Tyack & Cuban, 1995). Because the CCSS-M were intentionally constructed to allow flexibility in implementation across districts and schools, we can expect that the extent to which new standards led to improved student achievement varied widely depending on district implementation strategies. After reviewing the literature on CCSS-M implementation, we identify teacher professional learning (PL) around instructional practices as a particularly promising standards implementation strategy. We then examine changes in student outcomes across schools in Chicago based on their use of

supports for PL around the standards, asking the following questions:

1. How did students' reports of instructional practices, their test scores, grades, and pass rates in math change with implementation of the CCSS-M, based on their schools' level of participation in standards-related PL?
2. How did changes in math outcomes differ for students with low, average, and high prior math achievement?

The CCSS-M Transition: Research on Implementation and Student Outcomes

The CCSS-M contain two different types of standards. Content standards describe what students should learn at each grade level, whereas practice standards call for learning “processes and proficiencies” to support conceptual understanding across content areas. The practice standards include asking students to “make sense of problems and persevere in solving them,” “reason abstractly and quantitatively,” and “construct viable arguments and critique the reasoning of others” (Common Core State Standards Initiative, 2020c). Conceptual understanding, widely understood to derive from knowledge of key mathematical principles and how they relate to one another (Hiebert &



Lefevre, 1986; Rittle-Johnson et al., 2001), can be fostered by having students engage in these types of practices (Camburn & Han, 2011; Hattie, 2009; Prawat, 1989; Rittle-Johnson & Schneider, 2015). Integrated with instruction on procedural learning, they have been shown to improve student achievement (Baroody, 2003; Fennema et al., 1996; Kazemi, 1998; Newmann et al., 2001; Rittle-Johnson et al., 2001; Wenglinsky, 2002).

Research on whether the CCSS-M have actually led to improvements in math achievement is sparse and filled with caveats. Polikoff (2017) identified four priority areas for research around the CCSS-M: (1) studies on the effects of the standards on student outcomes, (2) research on implementation, (3) research on how supports provided through district and school implementation plans change teacher practice and student outcomes, and (4) examination of heterogeneous effects on subgroups of students. To date, there are only a few studies in each of these areas, and many gaps remain in what is known.

Studies on the Effects of Standards on Student Outcomes

Because the standards were adopted by almost all states at the same time, it is difficult to find comparison groups to develop strong causal estimates of their effects on students. Several studies have used states as the unit of analysis, comparing student achievement in years before and after standards implementation. One study compared states that were early adopters with late-adopting states, among states with low-rigor standards prior to the Common Core, and it found moderate effects on fourth-grade math scores (around 0.10 *SD*), only modest effects on eighth-grade math scores (around 0.04 *SD*), and no effects among economically disadvantaged students (Bleiberg, 2020). Another study found small improvements in ACT (American College Test) scores in Kentucky (0.03–0.04 *SD*), comparing cohorts of students exposed to the standards with earlier cohorts, but it was unable to conclude whether the impacts were due to the standards or other changes in the state (Xu & Cepa, 2018). A series of studies compared changes in CCSS adopter states with strong implementation with a very small number (four to eight) of nonadopter states and found a mix of null, very small positive, and very small negative effects on student achievement, based on different specifications (Loveless 2014, 2016). A further study compared states based on the degree to which the new standards differed from their previous standards and found slightly negative long-term impacts of the CCSS-M that grew more negative over time (–0.10 to –0.06 *SD*; Song et al., 2019). One further study did consider local variation, comparing districts within the same state, based on their adoption of CCSS-aligned textbooks. It found that implementing districts had a slightly higher percentage of middle school students scoring proficient on the state achievement test compared with nonimplementing districts,

while differences in math achievement across elementary students were insignificant (Gao & Lafortune, 2019).

Thus, when examining changes in achievement, the findings are not consistent, the existing research typically examines statewide averages without considering variation across schools and districts, and it often focuses on changes in the content standards. Moreover, existing research uses standardized test scores as the sole measure of achievement. State tests may not be adequate for measuring student learning from standards-aligned instruction; Polikoff and Porter (2014) found surprisingly small associations between standards-aligned instruction and student test scores. Besides looking at test scores, research has yet to consider whether students' course grades changed along with their probability of passing math. Students' grades in their courses have been shown to be stronger predictors than test scores of high school success, graduation, college enrollment, and college persistence (Allensworth & Clark, 2020; Easton et al., 2017). Course failures are critical math outcomes, given the demonstrated risk of increased failure rates in response to higher academic demand (Allensworth et al., 2009; Clotfelter et al., 2015; Simzar et al., 2016).

Research on Standards Implementation

At the outset of CCSS-M implementation, administrators expressed concerns about teacher preparation to teach math in an entirely new way (McLaughlin et al., 2014; Rentner & Kober, 2014). Studies of CCSS-M implementation across standards-adopting states have found that teachers report making significant changes to their instruction in response to the standards (Bay-Williams et al., 2016; Kane et al., 2016; Opfer et al., 2016). However, studies have also shown considerable variation in the extent to which teachers have changed their instruction (Edgerton & Desimone, 2018), with those who teach higher proportions of low-achieving students spending less instructional time engaged in standards-aligned practices than others (Schweig, Kaufman, & Opfer, 2020). Whether students themselves perceive differences in the types of instructional tasks they engage in during their math classes has not yet been studied, and instructional changes may be perceived differently for different subgroups of students.

Research on the Effects of Implementation Supports on Student Outcomes

While the CCSS-M were adopted by states, how to implement them in schools has been largely left up to districts (Desimone et al., 2019). Districts could support instructional change through many different mechanisms, including adopting new textbooks, digital tools, and other curricular resources; making efforts to improve teachers' math skills; aligning the scope and sequence of instruction in different

classes; and investing in professional development (PD) to support new instructional practices, PL communities, or coaching (Desimone et al., 2019; Kane et al., 2016). Studies of materials, curriculum, and textbooks have not found significant effects on student outcomes from changing to a new textbook in response to standards, or using a curriculum that was designed pre- versus post-CCSS-M, although a few particular curricula seem to be more effective than others (Blazar et al., 2019; Kane et al., 2016). However, there is evidence that more PL around the standards is related to stronger student outcomes. Kane et al. (2016) found that students scored higher on CCSS-aligned math assessments in schools where teachers had more PD around the standards and higher percentages of teachers who felt knowledgeable about them, controlling for differences in prior-year student achievement and teacher effectiveness.

PD has long been considered a primary lever for enacting instructional change (Desimone, 2009; Kennedy, 2016; Opfer & Pedder, 2011). Greater investments in PD provide legitimacy and consistency to what is being asked of teachers as well as the opportunity to learn new content and skills (Desimone, 2002, 2009; Desimone et al., 2019; Desimone & Garet, 2015). As states transitioned to the new standards, there were major challenges in providing sufficient high-quality PD (Kober et al., 2013; Rentner & Kober, 2014). Edgerton (2020) found that many districts lacked the capacity to provide ongoing support to teachers. Several years after standards-adopting states began implementing the CCSS-M, only about a third of teachers reported feeling like they had high levels of preparation to teach the new standards (Kane et al., 2016; Scholastic & Bill & Melinda Gates Foundation, 2014). Teachers felt particularly in need of PL on differentiating instruction for students with different levels of prior achievement and helping students achieve the practice standards (Hamilton et al., 2016). Variation in the use of PL across schools and districts could be a key factor for understanding variation in student outcomes with CCSS-M implementation.

Research on Instruction for Heterogeneous Students Under the CCSS-M

While the CCSS guidelines recognize that students with disabilities, English learners (ELs), and low-achieving students may need special supports to meet the expectations of the new standards (Common Core State Standards Initiative, 2020a, 2020b), they leave the details of how to provide these accommodations up to teachers, schools, and districts. In particular, a number of scholars have pointed out the need for special attention to the ways students with a track record of underperformance in math have experienced the new standards (Haager & Vaughn, 2013; Powell et al., 2013; Saunders et al., 2013).

Previous research suggests that, absent effective supports, students with a history of low achievement in math may benefit less than other students or even have worse academic outcomes in response to reforms that increase rigor in their math classes (Allensworth et al., 2009; Clotfelter et al., 2015; Gamoran & Hannigan, 2000; Simzar et al., 2016). Students with a history of low achievement may become disengaged and stop trying when academic rigor increases without simultaneous efforts to help them meet the new demands (Bandura, 1986; Bouffard-Bouchard, 1990; Lent et al., 1984; Shernoff et al., 2003). One of the most frequent concerns educators have expressed about the CCSS-M is lack of confidence in their ability to help students with low prior achievement attain the skills expected by the new standards (Haager & Vaughn, 2013; Hamilton et al., 2016; Powell et al., 2013; Saunders et al., 2013).

At the same time, the standards were adopted to ensure that all students have access to a rigorous curriculum and instructional practices that promote a deep understanding of mathematics. Low-achieving students frequently have been grouped in classes that focus heavily on procedures and computation (National Council of Teachers of Mathematics, 2018). Divorced from instruction to support comprehension of key mathematical concepts, this type of instruction has been shown to make students less motivated to discern the reasoning behind the procedures they use (Boaler et al., 2000; Hiebert & Lefevre, 1986). Over time, students become less equipped to undertake rigorous math problems, leading to the observed widening of gaps between low and high achievers over time (Jordan et al., 2009; Niemi, 1996; Richland et al., 2012). For low-achieving students with low conceptual understanding of math, instruction to support these practice standards may be especially critical.

How This Study Contributes

In this study, we examine changes in students' instructional experiences and achievement after the adoption of new standards, based on the extensiveness of teachers' PL around the standards across schools in Chicago. We examine a wider variety of student outcomes than in existing research on CCSS-M, including test scores, grades, rates of course passing, and student reports of engaging in standards-aligned instructional practices. After examining average outcomes of the CCSS-M implementation across students, we disaggregate the analyses for students with low, average, and high prior-year achievement. The "low-achieving" group includes most students with disabilities at each grade, as well as students facing other obstacles, broadly conceived, that contribute to below-grade level math achievement. Arguably, this group contains those students whom standards-based reforms aim to help the most, but who previous studies suggest will be most likely

to struggle with the increased academic demand of the new math standards.

Like the Kane et al. (2016) study, we use variation in participation in PL around the CCSS-M to examine the relationship between PD and changes in student outcomes. We also incorporate a longitudinal component, comparing differences in outcomes in poststandards years relative to prestandards years, similar to Song et al. (2019), Loveless (2014, 2016), and Bleiberg (2020). As with those studies, and with the Kane et al. (2016) study, causal inferences are limited because the schools with different levels of implementation are not random, just as the states with weaker standards pre-CCSS-M were not random in the studies of states. We include school fixed effects in the models so as to compare schools with themselves in prior years; this provides more confidence that changes in outcomes across schools with different levels of participation in PL were not simply a result of initial differences between schools.

We are able to do this work using detailed data on students in the middle grades (Grades 6 to 8) from the Chicago Public Schools (CPS; n.d.), a large urban district with about 500 schools that implemented the CCSS-M through a teacher leader model that provided PD and resources to encourage instructional change around the CCSS-M. The decentralized organizational structure of schools in the district resulted in significant variation in schools' participation in the district's PD standards implementation strategy. We leverage this variation to examine the relationship between PD on the CCSS-M and changes in student outcomes.

Standards Implementation in Chicago Public Schools: Promoting High-Quality Instruction Through Teacher Leaders

After Illinois adopted the CCSS-M in 2010, the CPS Department of Mathematics launched a multiyear effort, beginning in 2013, to support teachers' transition to the new standards. Their focus was on helping teachers modify their instructional practices so that students would develop the procedural and conceptual understanding called for by the practice standards. As described by one district leader to a member of the study team in an interview for an associated qualitative study, "this [strategy] was going to be about high-quality instruction. Then if the materials changed, the standards changed, if anything, the assessment changes, it doesn't matter because this is still good teaching." The district partnered with local universities to develop and provide content for PD sessions, which was informed by Charlotte Danielson's Framework for Teaching (Danielson, 2009) and the Teaching for Robust Understanding (TRU) Framework (Schoenfeld & Teaching for Robust Understanding Project, 2016), and it emphasized creating equitable, student-centered learning environments.

A primary feature of their PL program was a teacher leader model. Two or three teachers from each school were designated teacher leaders and attended Teacher Leader Institutes. In schools with higher proportions of low-achieving students, teacher leaders and their administrators were offered the opportunity to participate in math team learning communities where they could develop school-based PL plans and share across schools. All teacher leaders were expected to develop their own classroom practice to better align with the goals of the CCSS-M and support colleagues in their schools to make instructional improvements. The district offered a number of voluntary standards-focused workshops and opportunities for PD and collaboration that teachers could attend if they were not designated teacher leaders. The district also provided resources to support instruction and a list of recommended core curricula through a website called the "Knowledge Center." PD sessions often helped teachers use the materials provided by the Knowledge Center and the resources it contained to support PL within schools.

CPS schools differed considerably in the degree to which their teachers participated in PL around the standards (Gwynne & Cowhy, 2017), similar to what has been observed in studies of other places (Desimone et al., 2019; Kane et al., 2016). This variation allows us to analyze the relationship between the extensiveness of standards-related PL and changes in math instruction and student learning outcomes.

Method

Sample and Data

Our analyses are based on CPS students who were in the sixth through eighth grades at any point from 2011–2012 to 2016–2017. Students are included if they had a math score on the state assessment from the previous school year and were enrolled in regular CPS schools, which include neighborhood schools, magnet schools, and selective schools, but exclude alternative schools for dropout recovery and special education schools. Charter schools are also omitted from the analyses presented in the main text because we lacked their data on course grades and course passing rates.

Survey Data. A number of our variables come from surveys of teachers and students administered in the spring of each year in CPS. Participating in surveys is optional for individuals. However, schools receive reports on a subset of measures that are used for school improvement planning and district accountability. Thus, participation is strongly encouraged and time is usually set aside for teachers and students to take the surveys during the school day. Many of the survey measures show strong evidence of validity in their relationships with growth in student achievement over time

(Klugman et al., 2015; Nomi & Allensworth, 2013). As a result of the strong district support of the surveys, response rates are high. Teacher response rates in 2015 were 80.7%. Student survey response rates ranged from 74% to 83% for the surveys administered in the spring of 2011 through 2017.

Measure of Professional Learning. Our main independent variable, other than school year, is the extent to which teachers in each school participated in the district-wide PL program designed to support the new standards. Teachers' participation in PL is captured with the teacher survey administered in the spring of 2015, which was the first year (2014–2015) that all teachers in the district were expected to use standards-aligned instruction. Several questions asked teachers to describe their PD related to the new standards. See online Supplemental Appendix A for specific survey items. These were combined into a single measure with a reliability of 0.76 using a Rasch model (Wright & Masters, 1982). Since teacher responses on the survey are anonymous and cannot be tied to individual students and classrooms, we aggregated their responses to the school level to create a schoolwide measure of standards-related PD. Using hierarchical linear models, we calculated empirical Bayes estimates for each school, using a measurement model at Level 1, with teachers nested within schools at Levels 2 and 3. The school-level reliability of the measure was 0.72.

Using the empirical Bayes estimates from the hierarchical linear models, schools were categorized into three equal groups, henceforth identified as schools with extensive, average, or limited standards-related PL. The categories were determined including schools with middle grades and high schools so that parallel work on high schools would be comparable; this results in uneven numbers of schools in the categories for this study. High school teachers participated less extensively in PL around the standards than teachers in the middle grades. The groups reflect variation in participation across schools in PL through the district's teacher leader model, variation in any additional PD teachers experienced through other PL opportunities, and variation in school-based PD sessions run by teacher leaders. Responses to survey items across the three implementation groups are provided in online Supplemental Appendix B.

Schools with more extensive PL had slightly lower achievement, served students from neighborhoods with higher levels of poverty, and served more Black students (see Table 1). This corresponds with district efforts to offer additional PD supports for schools with higher proportions of low-achieving students. At the same time, the differences in mean achievement and student characteristics between the three groups are fairly small, relative to the variation within the groups. All three groups of schools are heterogeneous along all dimensions of student body composition, with considerable overlap across them on all student characteristics. Moreover, all of the statistical models control for

differences in student composition, and the use of school fixed effects provides additional confidence that the changes we observe in outcomes across schools in each PL category were not a result of initial differences between schools.

Student-Level Variables. District administrative records include information about students' background characteristics (race/ethnicity, gender, special education status, and census block group) for all students enrolled in the district each year. To control for students' socioeconomic status, we use variables created from U.S. Census data measuring neighborhood poverty (percentage of families under the poverty line and percentage of males unemployed) and neighborhood social status (median family income and average education levels), linked by their residential census block group. These allow a more fine-tuned measure of family background than an indicator of free and reduced lunch status, since more than 80% of students in CPS qualify. The Census measures are standardized across all census block groups in the city, not just the blocks in which CPS students live.

Student Reports of Math Practices. For each survey administration, students are asked questions specifically about practices they do in their math class (Math Instructional Practices). These practices correspond with the types of activities that are aligned with the CCSS-M Standards for Mathematical Practice and are consistent with the practices emphasized in district-sponsored PD around the standards. The questions were combined into a Rasch measure with a reliability of 0.57. See online Supplemental Appendix A for the specific survey questions.

Student Achievement. We use student transcript data for information on math course grades and whether students passed their math course. Course grades are available by semester, and some students take multiple math classes. We take the average for each student of their grades in math (measured on a 0- to 4-point scale) in each semester class. We calculated their math course pass rate based on whether they passed (1) or failed (0) each semester of each math course they took.

During the period covered by this study (2011–2017), CPS students took a number of different standardized tests. The old state test (ISAT [Illinois Standards Achievement Test]) was administered through the spring of 2014. The PARCC [Partnership for Assessment of Readiness for College and Careers], which was developed as an assessment for the CCSS-M, was administered beginning in the spring of 2015. Thus, the change in the state test occurred in the same year as standards implementation. Another math assessment, the NWEA-MAP (Northwest Evaluation Association Measures of Academic Progress), was given to all third- through eighth-grade students in every year from

TABLE 1
Characteristics of Schools in the Three Professional Learning Groups

Professional learning in 2015	Number of schools	Characteristic	<i>M</i>	<i>SD</i>	Min	Max
Limited	110	Prior-year average math score	0.09	0.61	-0.84	2.32
		Neighborhood poverty	0.22	0.56	-1.69	1.43
		Percent White	10%	17%	0%	67%
		Percent Black	41%	40%	0%	100%
		Percent Latino	43%	38%	0%	99%
		Percent special education	14%	7%	0%	34%
		Percent taking algebra	6%	9%	0%	39%
		Professional learning around the standards	-1.34	0.57	-3.06	-0.64
Average	172	Prior-year average math score	-0.04	0.46	-0.80	1.63
		Neighborhood poverty	0.31	0.60	-1.25	1.43
		Percent White	9%	18%	0%	73%
		Percent Black	47%	44%	0%	100%
		Percent Latino	40%	39%	0%	100%
		Percent special education	15%	6%	0%	34%
		Percent taking algebra	6%	9%	0%	49%
		Professional learning around the standards	-0.16	0.25	-0.64	0.26
Extensive	186	Prior-year average math score	-0.15	0.47	-1.07	2.02
		Neighborhood poverty	0.48	0.54	-1.09	1.52
		Percent White	5%	12%	0%	81%
		Percent Black	58%	44%	0%	100%
		Percent Latino	34%	39%	0%	100%
		Percent special education	15%	6%	0%	38%
		Percent taking algebra	4%	7%	0%	35%
		Professional learning around the standards	0.94	0.54	0.26	3.02

Note. Means are calculated based on the student population in the 2014–2015 school year.

2013 onward. This exam has more precision than the state tests because it is computer adaptive. It also correlates with the ISAT and PARCC at between 0.86 and 0.90, depending on grade and year, disattenuated for measurement error, based on our analysis of CPS data. Because the NWEA was given in years before and after standards implementation, we use NWEA scores as our outcome measure of math skills. However, so that we have a prescore measure for the year 2013, we use students' performance on the ISAT or PARCC from the prior year (2011 through 2016) as a control variable for prior achievement, standardized by grade and year. Online Supplemental Appendix C provides an analysis of change in PARCC scores over the post-CCSSM years by the extensiveness of PL in the school.

Table 2 provides details on variables for those students who have data on all outcomes, which is the sample used for the analyses presented in the main text, as well as the full data set. We limit the main analysis to students with complete data so the results are comparable across outcomes; the same students are included for each analysis. However, this substantially reduces the analytic sample down from the

population of students in the district. We have test score data for almost all students in the district, but students who transferred midyear do not have spring test data. We have transcript data for all students except those enrolled at charter schools. Survey response rates are about 80%. By including only students with data on all four outcomes, the analytic sample consists of 213,965 students in 418 schools versus 325,275 students in 478 schools. The characteristics of the reduced and full samples are very similar. We ran models for each outcome with the entire population of students who have data on the outcome to confirm that our conclusions would be the same, as described in online Supplemental Appendix C.

To answer the second research question, we divide the data based on students' prior-year achievement levels. Students in the low-achievement group scored at least a half standard deviation below the district average for their grade level on the prior-year state assessment, those in the high-achievement group had scores at least half a standard deviation above average, and the remainder were in the average achievement group. There are several differences other than

TABLE 2
Descriptive Statistics on Students in All Years (2013–2017)

Variable type	Variable name	Students with complete data		Students with any data		Students with low initial achievement		Students with average initial achievement		Students with high initial achievement	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Outcomes	Instructional practices	0.00	1.00	-0.01	1.01	-0.08	1.09	0.02	1.00	0.05	0.89
	NWEA-MAP score	0.00	1.00	-0.03	1.01	-0.90	0.78	-0.03	0.61	0.93	0.71
	GPA	2.58	1.05	2.55	1.06	1.93	0.96	2.51	0.93	3.32	0.79
	Math pass rate	0.96	0.21	0.95	0.21	0.90	0.30	0.97	0.18	0.99	0.07
Student covariates	Concentrated poverty	0.19	0.81	0.23	0.80	0.36	0.76	0.22	0.78	-0.02	0.84
	Social status	-0.45	0.90	-0.46	0.88	-0.58	0.85	-0.51	0.87	-0.26	0.94
	Prior math score	0.05	1.00	0.01	1.00	-1.05	0.42	-0.01	0.28	1.22	0.62
	Grade 7	33%		33%		32%		34%		33%	
	Grade 8	33%		33%		32%		34%		33%	
	Black	33%		37%		44%		34%		22%	
	White	10%		9%		4%		7%		19%	
	Asian	4%		4%		1%		3%		9%	
	Latinx	51%		50%		50%		55%		48%	
	Other race or ethnicity	1%		1%		1%		1%		2%	
	Male	49%		50%		52%		47%		48%	
	Identified disability	12%		13%		29%		7%		2%	
	Enrolled in algebra	7%		6%		1%		5%		17%	
	Other types of math	0%		0%		0%		0%		1%	
Multiple types of math	3%		3%		1%		3%		6%		
School level	Limited PL in 2015	22%		24%		20%		21%		25%	
	Average PL in 2015	39%		38%		38%		39%		41%	
	Extensive PL in 2015	39%		38%		42%		40%		34%	
	Total <i>n</i>	312,364		442,013		95,648		119,805		96,911	

Note. Low initial achievement is less than 0.5 *SDs* below the district average on the prior-year state assessment in math for student grade level; high initial achievement is at least 0.5 *SD* above; average initial achievement is within a half standard deviation. Instructional practices and NWEA scores are standardized around the 2013 mean and standard deviation. NWEA-MAP = Northwest Evaluation Association Measures of Academic Progress; GPA = grade point average; PL = professional learning.

their prior-year math scores (see Table 2). On average, students in the public schools in Chicago tend to live in areas of the city with higher poverty rates than the city as a whole. However, students with high prior-year achievement come from census blocks with poverty levels that are at the city average. These students may be more representative of state or national population averages in terms of students' economic status than the other groups. Students in the group with low prior-year achievement are much more likely to have identified disabilities (29% vs. 2% in the high-achievement group).

Analytic Methods

We take a difference-in-differences (DiD) approach, comparing changes in student outcomes in postimplementation years to preimplementation years, in schools with extensive and moderate implementation of PL around the standards to those with limited PL. All schools were expected to

implement the standards beginning in the 2014–2015 school year. This is also the year in which we measured teachers' participation in PL around the standards. We refer to 2015 as the year when the standards and the PL around the standards were implemented, with 2011 to 2014 as preimplementation years and 2015 to 2017 as postimplementation years.

The first difference that we model is based on school year. We compare outcomes in the spring of 2015, 2016, and 2017 (postimplementation years) with outcomes in the spring of 2014 to discern whether student outcomes were higher in postimplementation years. Each year is modeled as its own dummy variable, with 2014 as the excluded group, so that we can observe whether any changes were short-lived, sustained, or grew over time. If the CCSS-M brought changes in instruction and student outcomes, we should see that outcomes are different in 2015 and later years, relative to 2014. We include a series of student covariates so that we compare the outcomes of students with the same backgrounds over time.

Because other initiatives were occurring in the district at the same time that it adopted the CCSS-M, we cannot fully attribute any of the changes observed in student outcomes in postimplementation years to standards implementation. Therefore, we focus on teacher participation in PL around the standards in the 2014–2015 school year as one aspect of standards implementation, discerning whether schools with extensive or average participation in PL saw greater changes in student outcomes than schools with limited PL. Because teacher participation in PL was minimal in schools with limited PL, the changes we observe postimplementation for schools with limited PL mostly reflect other changes happening in the district. Thus, the differences estimated for these schools minus the differences observed in the other schools with different levels of PL reflect the effects of participation in PL around the standards. This approach provides more confidence that observed changes were due to a component of standards implementation rather than other changes occurring in the district.

One concern with this approach is that the schools that participate more fully in PL around the standards are different in some unmeasured ways from schools with limited PL, and that these differences might also be related to student outcomes. To mitigate this concern, we include school fixed effects that capture stable characteristics of schools. Thus, we are comparing school outcomes in postimplementation years with outcomes in the same schools preimplementation—comparing schools with themselves. We include dummy variables for each school year with 2014 as the excluded group, alone and interacted with PL. If the groups are comparable, with the covariates and school fixed effects, we should see no significant difference in the years before implementation, relative to 2014 (the year before implementation). We should see differences in 2015 and later years, relative to 2014, if PL in 2015 was related to improvements in student achievement. The model is as follows:

$$Y_{ist} = \sum_{t=1}^6 \beta_t \text{year}_t + \sum_{c=7}^{18} \gamma_c \text{year}_t * I_s + \alpha_s + \Omega Z_{ist} + \varepsilon_{ist}$$

The outcome Y for student i in school s in time t , where

- β_1 to β_6 capture the differences in the outcome in each year compared with 2014 through a series of dummy variables for time (D2011–D2017, excluding 2014);
- I_s are dummy variables for two of the implementation groups—schools with average PL and schools with extensive PL;
- γ_c are the DiD coefficients for schools with different levels on the intervention, I (either extensive or average implementation). Among schools with limited PL (the excluded group), the difference in the outcome from 2014 to 2015 is the main effect for D2015, β_4 , while for schools with average PL, the difference

is β_4 plus the coefficient on the interaction of D2015

* ImpAve, γ_{10} .

α_s are school fixed effects;

Z_{ist} are student-specific covariates; and

standard errors were clustered by school.

We also ran a model that used a continuous variable of PL, rather than categories of limited, average, and extensive PL (see online Supplemental Tables C4 and C5 in Appendix C). We highlight the categorical models since they are used to make a figure and table that are easier to interpret.

We then run the same models separately for each of the three subgroups defined by students' achievement in the prior year to discern whether there were differences in outcomes for students with different prior-year achievement levels. By running the models separately, the covariates are specific to each of the subgroups, and the estimates are not influenced by outcomes for students in the other achievement groups.

Limitations

There are a number of caveats to this study. First, the schools with extensive PL were not a random group. Administrators could elect to have their schools participate in standards-related PL, and teachers could elect to participate. Those who participated more fully may have been more receptive to PL. Online Supplemental Appendix B provides more information on the three implementation groups. Each group contained a heterogeneous mix of schools, but schools with more extensive participation had somewhat lower achievement, on average. Their teachers might have felt a greater press for change. Requiring extensive PL in schools with less buy-in might produce different results than shown here.

Second, the magnitude of change due to PL may be an underestimate of the total effects because we only examine differences relative to other schools—all of whom received some degree of PL around the standards. As shown in online Supplemental Appendix B, not all teachers in schools with extensive PL participated in standards-related PD, while a small percentage of teachers in schools with limited PL participated regularly in PL. The relationships we measure are further diminished by measurement error that exists in our school-level measure of teachers' participation in PL.

Third, these changes are occurring within a context of other policy changes in the district, including a new teacher evaluation system, and restrictions on the use of suspensions coupled with more restorative discipline strategies. While we leverage differences in participation in standards-based PL to tease out the influence of standards-based support, it could be that standards-based PL was effective only because these other policy elements were also in place, or were in some way due to these other district policies. Fourth, these

methods isolate one particular element of implementation (PL) and do not identify the total influence of the standards. There is not a way to tease out the results of the complete district implementation plan from other district-wide changes. These are also conservative estimates of the influence of PL around the standards, as we measure participation in PD in the first year of district-wide standards implementation, while student outcomes are measured through 3 years of implementation. We do this so the treatment variable (teacher reports of participation in PD in the 2014–2015 school year) occurs prior to the time any outcome variables would be observed (spring 2014–2015 through 2016–2017).

Results

Coefficients from the first set of models are shown in Table 3. The year fixed effects in the first several rows show how students' outcomes changed in each year, relative to 2014, among schools with limited PL around the standards. There were improvements in most outcomes over time, but we cannot determine whether they were due to the standards, or to other district initiatives. The next two sets of coefficients include interaction terms between year and whether schools engaged in average or extensive levels of PL around the standards. These coefficients show the extent to which outcomes changed more in schools with average or extensive PL around the standards than in schools with limited PL. These coefficients should be added to the main year fixed effects to determine the overall change in the outcome for schools with average or extensive PL around the standards, relative to 2014. To make this easier to see, Figure 1 graphs these changes over time. The 2011 to 2014 values in the figures show how much the schools started off in different places on the outcomes, but the school fixed effects control for any differences when estimating the DiD coefficients in the models.

In general, schools with extensive PL around the standards in 2015 showed significantly more growth in student outcomes than schools with limited PL in 2015 and later years, while schools with average PL showed few differences from schools with limited PL. Neither group showed significantly different growth compared with schools with limited PL in 2011 to 2014, the years before implementation. Starting in 2015, students in schools with extensive PL reported significantly greater improvement in Math Instructional Practices than students in schools with limited PL around the standards (0.080 *SD*). The differences relative to schools with limited PL were sustained over time. Students in schools with average levels of PL also reported significantly larger improvements in instructional practice than schools with limited PL, but these differences were only significant in 2016.

Student academic outcomes also improved more in schools with extensive standards-related PL than those with

limited PL. Math course grades improved by 0.087 GPA (grade point average) points more in schools with extensive PL than in schools with limited PL in 2015, with the difference growing to 0.130 GPA points by 2017. Schools with average standards-related PL did not show significantly larger improvements in GPA than schools with limited PL. Math course pass rates improved more in schools with both average and extensive PL compared with schools with limited PL, improving by 1 to 2 percentage points more in 2015, 2016, and 2017. NWEA scores grew more in schools with extensive PL than in schools with limited PL in 2016 and 2017 (0.061 *SD* in 2016 and 0.069 *SD* in 2017), but scores were not significantly higher in schools with average PL than in schools with limited PL. The differences in NWEA scores in schools with extensive PL are similar to impacts on PARCC scores found by Kane et al. (2016), who found that an additional 2 days of PL each year improved PARCC scores by 0.045 *SD*.

Differences by Students' Prior Achievement

Looking at all students together obscures differences based on students' initial achievement levels. Schools with extensive PL showed significantly larger improvements in achievement than schools with limited PL among their low- and average-achieving students, but there were few significant differences in the achievement of their high-achieving students (see Table 4). For students with low initial achievement, math GPAs improved by 0.161 points more in 2017, and pass rates improved by 3.1 percentage points more, if they were in schools with extensive PL than in schools with limited PL. Low-achieving students' NWEA scores improved by 0.078 *SD* more in schools with extensive PL than in schools with limited PL in 2016, relative to scores in 2014. Low-achieving students also showed more improvements in grades and pass rates in schools with average PL around the standards than in schools with limited PL. At average-PL schools, low-achieving students' GPAs improved by more than 0.10 points in 2015 and 2017, and pass rates improved by about 3 additional percentage points in all post-implementation years, relative to grades and pass rates in 2014.

Average-achieving students also showed significantly greater improvements in test scores, math course grades, and pass rates in schools with extensive PL than in schools with limited PL, with test scores improving 0.093 *SD* more, grades by 0.151 points more, and pass rates by 1.4 percentage points more. Average-achieving students did not show significantly more improvements if in schools with average levels of PL, with the exception of showing higher pass rates in 2016 by 1.3 percentage points.

Table 5 shows the results as averages, rather than coefficients that need to be added together, and shows how the groups were different in the base year. The values for 2014 are the simple means for each subgroup, based on students'

TABLE 3
Coefficients From Models Predicting Outcomes With the Extensiveness of School Professional Learning

	Coefficient	Math practices (in <i>SD</i> units)		NWEA score (in <i>SD</i> units)		GPA (in GPA points)		Pass math (in percentage points)	
		Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Year fixed effects (relative to 2014)									
	2011 (Pre-CCSSM)	-0.103	0.032**			-0.202	0.038****	-3.2%	0.7%****
	2012 (Pre-CCSSM)	-0.058	0.030			-0.081	0.036*	-1.1%	0.5%*
	2013 (Pre-CCSSM)	0.035	0.027	-0.070	0.018****	-0.033	0.022	-0.8%	0.3%**
	2015 (Post)	0.083	0.026**	0.082	0.013****	0.007	0.022	-0.1%	0.4%
	2016 (Post)	0.096	0.025****	0.123	0.019****	0.088	0.027***	0.8%	0.4%*
	2017 (Post)	0.093	0.034**	0.141	0.024****	0.079	0.028**	0.8%	0.4%*
Schools with average PL (relative to low PL)									
	2011 × AveragePL (Pre)	0.007	0.042			0.112	0.046*	1.8%	0.8%*
	2012 × AveragePL (Pre)	0.013	0.038			0.021	0.044	0.0%	0.7%
	2013 × AveragePL (Pre)	-0.015	0.035	-0.023	0.025	0.027	0.029	0.5%	0.5%
	2015 × AveragePL (Post)	0.014	0.032	0.000	0.018	0.048	0.031	1.3%	0.5%*
	2016 × AveragePL (Post)	0.088	0.035*	0.018	0.024	0.028	0.036	1.7%	0.5%**
	2017 × AveragePL (Post)	0.070	0.044	0.018	0.031	0.064	0.038	1.4%	0.6%*
Schools with extensive PL (relative to low PL)									
	2011 × HighPL (Pre)	-0.049	0.039			0.067	0.046	0.3%	0.9%
	2012 × HighPL (Pre)	-0.048	0.037			0.039	0.043	-0.7%	0.8%
	2013 × HighPL (Pre)	-0.033	0.033	-0.025	0.023	0.009	0.032	-0.4%	0.5%
	2015 × HighPL (Post)	0.080	0.039*	0.015	0.021	0.087	0.031**	1.5%	0.5%**
	2016 × HighPL (Post)	0.060	0.033	0.061	0.027*	0.128	0.036***	2.0%	0.6%***
	2017 × HighPL (Post)	0.099	0.044*	0.069	0.033*	0.130	0.038***	1.7%	0.6%***
Covariates									
	Grade 7	-0.037	0.008****	0.309	0.010****	-0.013	0.014	-0.2%	0.2%
	Grade 8	-0.047	0.010****	0.563	0.010****	0.106	0.014****	2.3%	0.2%****
	Algebra	-0.021	0.014	0.215	0.014****	-0.124	0.019****	-2.0%	0.2%****
	Other math	-0.127	0.099	0.527	0.065****	0.164	0.113	-0.3%	1.7%
	Multiple math	0.055	0.020**	0.117	0.021****	0.159	0.031****	-0.1%	0.4%
	White	-0.111	0.011****	0.093	0.010****	0.230	0.013****	0.5%	0.3%
	Asian	-0.077	0.014****	0.160	0.013****	0.369	0.021****	0.6%	0.4%
	Hispanic	-0.130	0.008****	0.027	0.007****	0.111	0.010****	0.8%	0.2%****
	Other	-0.086	0.016****	0.071	0.012****	0.128	0.018****	0.2%	0.3%
	Male	-0.041	0.005****	0.040	0.003****	-0.308	0.006****	-3.2%	0.1%****
	Poverty	0.007	0.003*	-0.010	0.002****	-0.015	0.003****	-0.1%	0.1%
	Social status	0.002	0.004	0.003	0.003	-0.002	0.004	-0.1%	0.1%
	Prior math score	0.095	0.003****	0.735	0.004****	0.558	0.006****	3.6%	0.2%****
	Intercept	-0.121	0.015****	-0.203	0.011****	2.707	0.015****	98.1%	0.3%****
	R^2	0.052		0.714		0.379		0.078	

Note. Coefficients come from models that also include school fixed effects, with 312,364 student records over 7 years in 418 schools. Standard errors were clustered by school. Math practices and NWEA scores were standardized using means and standard deviations from 2014. Teacher professional learning (PL) around the CCSS-M was measured in the 2014–2015 school year, which was the same year all schools were to begin implementing the standards. $*p < .05$. $**p < .01$. $***p < .001$. $****p < .0001$.

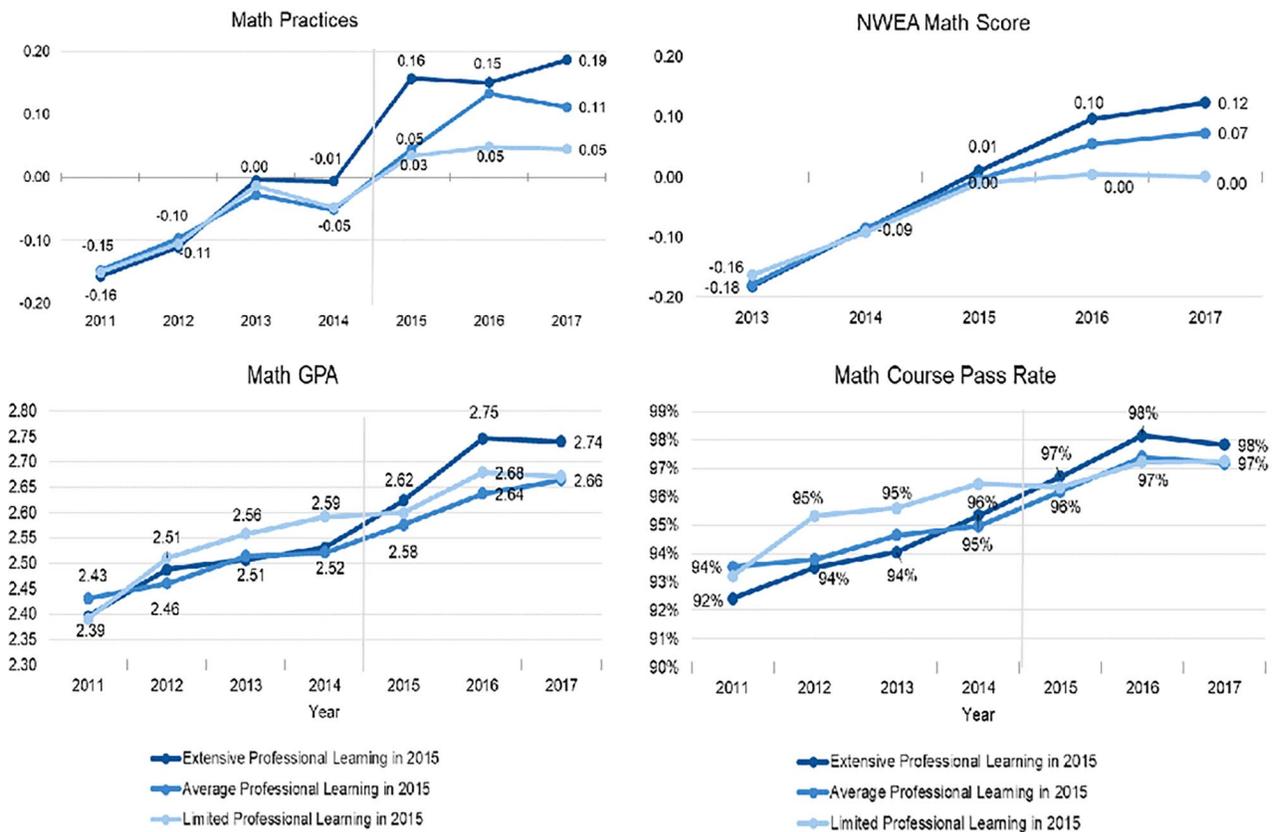


FIGURE 1. Student outcomes by the extensiveness of professional learning.

Note. Values for 2014 come from regression models that control for student race, ethnicity, gender, neighborhood poverty level, neighborhood social status, the type(s) of math class in which the student was enrolled, and math score on the state assessment in the prior spring. Changes relative to 2014 are based on coefficients from models with school and year fixed effects and the same covariates. Schools with average professional learning (PL) in 2015 showed significantly more growth than schools with limited PL in math practices in 2016 and pass rates in 2015, 2016, and 2017, relative to 2014. Schools with extensive PL in 2015 showed significantly more growth than schools with limited PL in math practices in 2015 and 2017, NWEA-MAP (Northwest Evaluation Association Measures of Academic Progress) scores in 2016 and 2017, GPAs (grade point averages) and pass rates in all years.

prior-year achievement, year, and school PL around the standards. The values for other years are calculated by adding the coefficients from Table 4 onto the 2014 subgroup means, with coefficients from interaction terms added to the year fixed effects for students in schools with average or extensive PL.

Prior to standards implementation, students with high prior-year achievement reported more frequently engaging in standards-aligned instructional practices than low-achieving students (by 0.23–0.25 *SD* in 2011 and 0.11–0.14 *SD* in 2014; see Table 5). In the postimplementation years, low-achieving students showed increases in standards-aligned practices of 0.15 *SD* in schools with limited PL and 0.24 *SD* in schools with extensive PL. Students with low achievement still reported less engagement in standards-aligned practices than students with high prior-year achievement in the same schools, but the differences were smaller, and low-achieving students in schools with extensive PL engaged in standards-aligned practices at higher rates than high-achieving students in schools with limited PL after standards implementation.

Overall, achievement was still lower among students who had lower prior achievement than other students, but in schools with extensive PL, the differences in GPA and pass rates between students with low and high test scores were narrower in postpolicy years than in prepolicy years. The difference in GPAs was 0.11 points smaller, and the difference in pass rates was 4 percentage points smaller. Differences in NWEA scores remained similar.

Alternative Specifications

Results from alternative models are presented in online Supplemental Appendix C; the conclusions are similar to those in the main text. Coefficients from models that use all cases with data on each outcome are included in online Supplemental Tables C1 to C3; the DiD coefficients in these models are larger and more consistently significant for predictions of math practices and NWEA scores among low- and average-achieving students. Coefficients from models that use a continuous measure of PL are shown in online Supplemental Tables C4 and C5; the patterns of which DiD

TABLE 4
Selected Coefficients Predicting Outcomes With School Professional Learning

Student Achievement Group	School PL	Year Fixed Effects (relative to 2014)	Math practices (in SD units)		NWEA score (in SD units)		GPA (in GPA points)		Pass rate (in percentage points)		
			Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	
Students with low prior-year achievement	Schools with limited PL (base)	2011 (Pre-CCSSM)	-0.179	0.041***			-0.212	0.050***	-6.2%	1.5%***	
		2012 (Pre-CCSSM)	-0.089	0.035*			-0.067	0.053	-2.4%	1.3%	
		2013 (Pre-CCSSM)	-0.040	0.039	-0.073	0.025**	-0.015	0.030	-1.6%	0.8%*	
		2015 (Post)	0.076	0.039	0.136	0.025***	0.027	0.034	0.0%	1.2%	
		2016 (Post)	0.081	0.042	0.164	0.028***	0.125	0.035***	2.2%	0.9%*	
		2017 (Post)	0.144	0.049**	0.182	0.031***	0.113	0.037**	2.5%	0.9%***	
		2011 × AveragePL (Pre)	0.026	0.051			0.132	0.061*	4.0%	1.9%*	
	Schools with average PL (relative to limited PL)	2012 × AveragePL (Pre)	0.018	0.045			0.030	0.063	1.0%	1.8%	
		2013 × AveragePL (Pre)	0.009	0.050	-0.046	0.035	0.019	0.043	0.9%	1.2%	
		2015 × AveragePL (Post)	0.021	0.048	-0.043	0.032	0.127	0.046**	3.7%	1.4%***	
		2016 × AveragePL (Post)	0.131	0.054*	0.005	0.037	0.094	0.049	3.7%	1.3%***	
		2017 × AveragePL (Post)	0.062	0.063	0.025	0.042	0.109	0.052*	3.0%	1.4%*	
		2011 × HighPL (Pre)	0.007	0.050			0.065	0.061	1.2%	1.9%	
		2012 × HighPL (Pre)	-0.047	0.044			0.025	0.062	-0.5%	1.7%	
	Students with average prior-year achievement	Schools with limited PL (base)	2013 × HighPL (Pre)	0.028	0.047	-0.056	0.032	-0.032	0.043	-1.2%	1.2%
			2015 × HighPL (Post)	0.134	0.061*	-0.007	0.033	0.145	0.047**	3.1%	1.4%*
			2016 × HighPL (Post)	0.117	0.051*	0.078	0.038*	0.168	0.048***	4.0%	1.3%***
2017 × HighPL (Post)			0.094	0.062	0.083	0.044	0.161	0.052**	3.1%	1.4%*	
2011 (Pre-CCSSM)			-0.099	0.040*			-0.222	0.056***	-3.2%	0.8%***	
2012 (Pre-CCSSM)			-0.061	0.036			-0.055	0.044	-0.6%	0.5%	
2013 (Pre-CCSSM)			0.038	0.032	-0.063	0.020**	-0.026	0.029	-0.9%	0.3%***	
2015 (Post)			0.092	0.035*	0.018	0.017	-0.006	0.032	-0.5%	0.4%	
2016 (Post)			0.090	0.035*	0.049	0.020*	0.065	0.037	0.1%	0.4%	
2017 (Post)			0.093	0.042*	0.016	0.028	0.025	0.037	0.1%	0.4%	
Students with average PL (relative to limited PL)	Schools with average PL (relative to limited PL)	2011 × AveragePL (Pre)	-0.010	0.050			0.138	0.065*	1.5%	0.9%	
		2012 × AveragePL (Pre)	0.020	0.045			-0.026	0.055	-1.2%	0.7%	
		2013 × AveragePL (Pre)	-0.006	0.042	-0.007	0.028	0.030	0.038	0.7%	0.5%	
		2015 × AveragePL (Post)	0.013	0.043	0.031	0.022	0.017	0.044	0.6%	0.6%	
		2016 × AveragePL (Post)	0.072	0.045	0.029	0.026	0.002	0.048	1.3%	0.6%*	
		2017 × AveragePL (Post)	0.063	0.052	0.031	0.035	0.078	0.050	1.2%	0.6%	
		2011 × HighPL (Pre)	-0.077	0.048			0.079	0.064	0.5%	0.9%	
	Schools with extensive PL (relative to limited PL)	2012 × HighPL (Pre)	-0.041	0.046			0.007	0.052	-1.4%	0.7%	
		2013 × HighPL (Pre)	-0.031	0.039	-0.028	0.027	0.008	0.041	-0.1%	0.5%	
		2015 × HighPL (Post)	0.057	0.049	0.039	0.024	0.068	0.042	1.1%	0.5%*	
2016 × HighPL (Post)	0.068	0.045	0.059	0.030*	0.134	0.048**	1.4%	0.6%*			
2017 × HighPL (Post)	0.116	0.052*	0.093	0.037*	0.151	0.049**	1.4%	0.6%*			

(continued)

TABLE 4 (CONTINUED)

Student Achievement Group	School PL	Year Fixed Effects (relative to 2014)	Math practices (in SD units)		NWEA score (in SD units)		GPA (in GPA points)		Pass rate (in percentage points)	
			Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Students with high prior-year achievement	Year fixed effects (relative to 2014)	2011 (Pre-CCSSM)	-0.042	0.039			-0.156	0.033***	-0.5%	0.2%*
		2012 (Pre-CCSSM)	-0.018	0.038			-0.103	0.032***	-0.3%	0.2%
		2013 (Pre-CCSSM)	0.097	0.036**	-0.067	0.023**	-0.047	0.021*	0.0%	0.1%
		2015 (Post)	0.068	0.029*	0.092	0.015***	-0.005	0.021	0.1%	0.1%
		2016 (Post)	0.108	0.029***	0.151	0.022***	0.073	0.028**	0.2%	0.1%
		2017 (Post)	0.050	0.038	0.207	0.025***	0.081	0.030**	0.3%	0.1%***
Schools with average PL	Schools with average PL	2011 × AveragePL (Pre)	0.020	0.053			0.075	0.041	0.2%	0.3%
		2012 × AveragePL (Pre)	-0.003	0.049			0.057	0.039	-0.1%	0.3%
		2013 × AveragePL (Pre)	-0.040	0.044	-0.027	0.027	0.035	0.030	0.0%	0.2%
		2015 × AveragePL (Post)	0.017	0.038	0.012	0.021	0.011	0.030	0.1%	0.2%
		2016 × AveragePL (Post)	0.078	0.040	0.030	0.028	-0.011	0.037	0.2%	0.2%
		2017 × AveragePL (Post)	0.090	0.050	0.020	0.034	0.012	0.040	0.1%	0.2%
Schools with extensive PL	Schools with extensive PL	2011 × HighPL (Pre)	-0.045	0.049			0.065	0.041	0.2%	0.3%
		2012 × HighPL (Pre)	-0.047	0.048			0.075	0.040	0.0%	0.3%
		2013 × HighPL (Pre)	-0.082	0.044	0.008	0.029	0.052	0.030	0.0%	0.2%
		2015 × HighPL (Post)	0.056	0.039	0.014	0.025	0.032	0.030	0.0%	0.2%
		2016 × HighPL (Post)	-0.008	0.040	0.044	0.031	0.051	0.037	0.1%	0.2%
		2017 × HighPL (Post)	0.072	0.051	0.061	0.037	0.076	0.040	0.0%	0.2%

Note. Models were run separately based on students' prior achievement levels. Models included school and year fixed effects, student race, ethnicity, gender, neighborhood poverty level, neighborhood social status, type(s) of math class, and prior spring math score, with 312,364 student records over 7 years in 418 schools. Standard errors were clustered by school. Low achievement ≥ 0.5 SD below the mean in the prior year; high achievement ≥ 0.5 SD above the mean. Math practices and NWEA scores were standardized using means and standard deviations from 2014. Coefficients on covariates are available in the online supplemental tables. Teacher professional learning (PL) around the CCSS-M was measured in the 2014–2015 school year, which was the same year all schools were to begin implementing the standards. NWEA = Northwest Evaluation Association; GPA = grade point average; CCSSM = Common Core State Standards in Mathematics; PL = professional learning.

* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE 5

Average Outcomes by Prior-Year Student Achievement, School Professional Learning Around the Standards and Year, Based on Model Estimates

Student achievement	School PL	Year	Math practices (in SD units)	NWEA score (in SD units)	GPA (in GPA points)	Pass math (in percentage points)
Students with low prior-year achievement	Schools with limited PL	2011 (Pre)	-0.30		1.75	87
		2012 (Pre)	-0.21		1.90	90
		2013 (Pre)	-0.17	-1.04	1.95	91
		2014 (Pre)	-0.13	-0.96	1.96	93
		2015	-0.05	-0.83	1.99	93
		2016	-0.04	-0.80	2.09	95
		2017	0.02	-0.78	2.08	95
	Schools with average PL	2011 (Pre)	-0.27		1.78	87
		2012 (Pre)	-0.19		1.83	87
		2013 (Pre)	-0.15	-1.06	1.87	88
		2014 (Pre)	-0.12	-0.94	1.86	89
		2015	-0.02	-0.85	2.02	93
		2016	0.10	-0.78	2.08	95
		2017	0.09	-0.74	2.08	94
	Schools with extensive PL	2011 (Pre)	-0.24		1.74	84
		2012 (Pre)	-0.20		1.84	87
		2013 (Pre)	-0.08	-1.07	1.84	87
		2014 (Pre)	-0.07	-0.94	1.89	90
		2015	0.14	-0.81	2.06	93
		2016	0.13	-0.70	2.18	96
		2017	0.17	-0.67	2.16	95
Students with average prior-year achievement	Schools with limited PL	2011 (Pre)	-0.14		2.34	95
		2012 (Pre)	-0.10		2.51	97
		2013 (Pre)	0.00	-0.11	2.54	97
		2014 (Pre)	-0.04	-0.05	2.56	98
		2015	0.05	-0.03	2.56	97
		2016	0.05	0.00	2.63	98
		2017	0.05	-0.03	2.59	98
	Schools with average PL	2011 (Pre)	-0.13		2.41	95
		2012 (Pre)	-0.06		2.42	95
		2013 (Pre)	0.01	-0.11	2.50	97
		2014 (Pre)	-0.02	-0.04	2.50	97
		2015	0.08	0.00	2.51	97
		2016	0.14	0.03	2.56	98
		2017	0.13	0.00	2.60	98
	Schools with extensive PL	2011 (Pre)	-0.15		2.32	94
		2012 (Pre)	-0.08		2.42	95
		2013 (Pre)	0.03	-0.14	2.45	96
		2014 (Pre)	0.03	-0.05	2.46	97
		2015	0.17	0.00	2.53	97
		2016	0.18	0.06	2.66	98
		2017	0.23	0.06	2.64	98
Students with high prior-year achievement	Schools with limited PL	2011 (Pre)	-0.05		3.22	99
		2012 (Pre)	-0.02		3.27	99
		2013 (Pre)	0.09	0.85	3.33	100
		2014 (Pre)	-0.01	0.91	3.38	100

(continued)

TABLE 5 (CONTINUED)

Student achievement	School PL	Year	Math practices (in <i>SD</i> units)	NWEA score (in <i>SD</i> units)	GPA (in GPA points)	Pass math (in percentage points)
		2015	0.06	1.00	3.37	100
		2016	0.10	1.06	3.45	100
		2017	0.04	1.12	3.46	100
	Schools with average PL	2011 (Pre)	-0.04		3.23	99
		2012 (Pre)	-0.04		3.26	99
		2013 (Pre)	0.04	0.77	3.29	99
		2014 (Pre)	-0.02	0.86	3.31	99
		2015	0.06	0.96	3.31	100
		2016	0.16	1.04	3.37	100
		2017	0.12	1.09	3.40	100
	Schools with extensive PL	2011 (Pre)	-0.01		3.18	99
		2012 (Pre)	0.01		3.24	99
		2013 (Pre)	0.09	0.73	3.28	99
		2014 (Pre)	0.07	0.79	3.27	99
		2015	0.20	0.90	3.30	100
		2016	0.17	0.99	3.40	100
		2017	0.20	1.06	3.43	100

Note. Values for 2014 are simple averages for each subgroup. Values for other years use coefficients from models controlling for school and year fixed effects, student race, ethnicity, gender, neighborhood poverty level, neighborhood social status, the type(s) of math class, and prior-year math score, added on to the 2014 values. NWEA = Northwest Evaluation Association; GPA = grade point average; CCSSM = Common Core State Standards in Mathematics; PL = professional learning.

coefficients are significant are similar to those of schools with extensive PL. Online Supplemental Table C6 provides an analysis of PARCC scores; these were significantly higher in 2016 for schools with average PL, and in both 2016 and 2017 in schools with extensive PL. Table C8 includes teacher fixed effects; this model has the potential to confound school PL and teacher effects, but there are significant relationships on a number of the DiD coefficients, and the size of the coefficients are similar to models without teacher fixed effects.

Discussion

While the CCSS identify a centralized set of knowledge and skills that students should master, strategies for implementing the new standards are intentionally left to districts and schools. This design has led to considerable variation in implementation, prompting calls for research on how district and school supports affect classroom practice and academic outcomes for various subgroups of students (Polikoff, 2017). In Chicago, the district developed a PL program that emphasized changing instructional practices in math through a teacher leader model. Evidence from this study suggests that the strategy seems to have paid off, increasing student engagement in instructional practices consistent with the new standards, reducing differences in

instructional experiences among students with different prior achievement levels, and showing improvements in test scores, grades, and pass rates in math classes. Achievement improved district-wide during this period for students of all achievement levels, but the improvements were greatest in schools with extensive participation in PL, particularly among students with low prior-year achievement. The patterns observed in Chicago are in line with the goals of the CCSS-M to improve instructional equity, closing achievement gaps in math course pass rates and grades.

The Benefits to Achievement Were Observed Most Consistently for Students With Low Achievement and Those at Risk of Failing Their Math Class: Why?

Study after study have shown that teachers worry about being able to teach to the new standards for low-achieving students, and they find it difficult to do so (Haager & Vaughn, 2013; Hamilton et al., 2016; Powell et al., 2013; Saunders et al., 2013). However, even if they find it difficult, the benefits may still be large for low-achieving students who were least likely to engage in practices that promote deep engagement and conceptual understanding prior to the standards. A prior study on double-dose algebra in Chicago also found improvements in math scores for

students with below-average math skills. In that case, as in the case of CCSS-M implementation, teachers received PD for how to do more student-centered instructional practices that support conceptual understanding, and the extra time gave them the flexibility to take the chance on teaching in new ways (Nomi & Allensworth, 2013). Changes in instructional practices may also have been less critical for the achievement of students who already were successful in math.

Chicago Provides an Example of Successful Implementation: Why Is It Different From Studies of CCSS-M Adopting States?

This analysis provides evidence of positive outcomes associated with CCSS-M implementation in one city. Because most studies analyze states as a whole, we don't know which other districts had successful implementation. There are a number of reasons why we might see positive results in Chicago when statewide results are null, or show improvements mostly for high-SES students—and these have implications for future efforts to improve achievement and equity in math.

First, it could be that CPS emphasized what ultimately matters most for student engagement and learning, particularly for low-achieving students. The district's support for new instructional practices meant that teachers were prepared to implement not just the content standards but also the practice standards. The Teaching for Robust Understanding and Danielson frameworks integrated into the PD around the standards further supported the emphasis on instructional practices, and the availability of instruction-ready resources for promoting student discussion and problem-solving around multiple solutions made it easier for teachers to try new techniques. The prior studies that found no positive effects, or positive effects of standards only on economically advantaged students (Bleiberg, 2020; Song et al., 2019), used change in the content standards as the independent variable; they could not speak to how much instructional practices changed in those states. It is possible that economically advantaged districts had more resources to invest in PD. It is also possible that raising content standards without improving the quality of instruction does not benefit low-achieving students; studies of prior policies intended to raise rigor in math classes have found null or negative effects on low-achieving students (Allensworth et al., 2009; Clotfelter et al., 2015; Gamoran & Hannigan, 2000; Simzar et al., 2016). Furthermore, the Polikoff and Porter (2014) study that found no relationship between teachers' instruction and student gains focused primarily on the content standards.

Second, the Chicago context is not representative of the state or nation as a whole; it is an 80% low-income district, where many students begin school with low achievement

relative to national averages. Significant achievement gains were only observed for students with low or average achievement relative to other students in Chicago. When results are averaged across all students, the benefits for students with low achievement may be obscured. Typical districts in the state and nation may be more similar to students who are high achieving in Chicago.

The teacher model was also run in a way that is quite different from PL approaches in other places. This might have influenced the success of schools with extensive PL. Rather than training teachers to implement curricula in a prescribed way, the district encouraged teachers to "try out" practices and share what worked with others. This approach—with teachers supporting each other around instructional change with access to expert knowledge and resources—requires experimenting and learning from failure. That can be frustrating for teachers and involves considerable uncertainty. But it is consistent with learning science and innovation research (Cobb & Jackson, 2011; Coburn et al., 2013; Coburn & Russell, 2008). It also mirrors the goals of the CCSS-M for students, where students are presented problems and given support to find different solutions themselves.

Prior studies also focused solely on test scores as outcomes. Improvements in test scores were less clear than improvements in grades or pass rates—test scores improved gradually over time while grades and pass rates improved in the implementation year and then continued to grow. If we had only looked at test scores, we would have added another study to the literature that shows ambiguous results. The standards themselves may not clearly translate into gains on standardized tests—or they may require considerable support for teachers to make instructional changes that have a positive impact and those changes may take time to be seen in test scores.

The Teacher Leader Model Was Successful for Some Schools but Not All Schools

The model seems to have been effective for schools that participated in standards-aligned PL most extensively. In these schools, improvements began in 2015 and continued to accumulate through 2017. This continued improvement may reflect continued participation in PL, which was offered by the district through 2017. It may also reflect the design of the teacher leader model, wherein multiple teacher leaders from each school worked together to improve instruction. This collaborative design could have improved internal capacity for changes that were sustainable over time.

At the same time, it is important not to lose sight of the wide variation in participation in PL across schools. Many schools participated in PL to only a limited degree, and the achievement gains for schools with average levels of PL were largely null relative to schools with limited PL, except among low-achieving students. Furthermore, this analysis

only shows what happened in schools serving the middle grades. Initially, we planned to include high schools in our analysis. However, only 8% of high schools (11 schools) engaged in extensive PL. The estimates of policy effects among high schools were not stable with different model specifications. This variation in schools' implementation suggests that the teacher leader model was not entirely sufficient for meeting the district's aim of reaching all its teachers.

Acknowledgments

We would like to thank Amy Cassata and Jeanne Century from UChicago STEM Education, University of Chicago, for their partnership and feedback, and sharing information about professional learning initiatives around the CCSS-M in the Chicago Public Schools. We would like to thank Jessica Mahon at the Chicago Public Schools for her partnership and input on this work. We are also grateful to Paloma Blandon and Arya Muralidharan for their assistance. The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A1600162 to the University of Chicago. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

ORCID iDs

Elaine Allensworth  <https://orcid.org/0000-0002-1795-5351>
Sarah Cashdollar  <https://orcid.org/0000-0002-8696-1659>

References

- Allensworth, E. M., & Clark, K. (2020). High school GPAs and ACT scores as predictors of college completion: Examining assumptions about consistency across high schools. *Educational Researcher*, 49(3), 198–211. <https://doi.org/10.3102/0013189X20902110>
- Allensworth, E. M., Nomi, T., Montgomery, N., & Lee, V. E. (2009). College preparatory curriculum for all: Academic consequences of requiring algebra and English I for ninth-graders in Chicago. *Educational Evaluation and Policy Analysis*, 31(4), 367–391. <https://doi.org/10.3102/0162373709343471>
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice Hall.
- Baroody, A. J. (2003). *The development of adaptive expertise and flexibility: The integration of conceptual and procedural knowledge*. Lawrence Erlbaum.
- Bay-Williams, J., Duffett, A., & Griffith, D. (2016). *Common Core Math in the K-8 classroom: Results from a national teacher survey*. Thomas B. Fordham Institute. <https://files.eric.ed.gov/fulltext/ED570138.pdf>
- Blazar, D., Heller, B., Kane, T., Polikoff, M., Staiger, D., Carrell, S., & Kurlaender, M. (2019). Learning by the book: Comparing math achievement growth by textbook in six Common Core states (Research Report). Center for Education Policy Research, Harvard University. https://cepr.harvard.edu/files/cepr/files/cepr-curriculum-report_learning-by-the-book.pdf
- Bleiberg, J. (2020). *Does the Common Core have a common effect? An exploration of effects on academically vulnerable students* (EdWorkingPaper: 20-213). <https://doi.org/10.26300/v0an-cz33>
- Boaler, J., William, D., & Brown, M. (2000). Students' experiences of ability grouping: Disaffection, polarisation and the construction of failure. *British Educational Research Journal*, 26(5), 631–648. <https://doi.org/10.1080/713651583>
- Bouffard-Bouchard, T. (1990). Influence of self-efficacy on performance in a cognitive task. *Journal of Social Psychology*, 130(3), 353–363. <https://doi.org/10.1080/00224545.1990.9924591>
- Camburn, E. M., & Han, S. W. (2011). Two decades of generalizable evidence on U.S. instruction from national surveys. *Teachers College Record*, 113(3), 561–610. <https://www.tcrecord.org>
- Chicago Public Schools. (n.d.). *Common Core Mathematics PD at CPS*. Chicago Public Schools.
- Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2015). The aftermath of accelerating algebra evidence from district policy initiatives. *Journal of Human Resources*, 50(1), 159–188. <https://doi.org/10.3368/jhr.50.1.159>
- Cobb, P., & Jackson, K. (2011). Towards an empirically grounded theory of action for improving the quality of mathematics teaching at scale. *Mathematics Teacher Education and Development*, 13(2), 6–33. <https://mtd.merga.net.au/index.php/mtd/article/view/44>
- Coburn, C. E., Mata, W. S., & Choi, L. (2013). The embeddedness of teachers' social networks: Evidence from a study of mathematics reform. *Sociology of Education*, 86(4), 311–342. <https://doi.org/10.1177/0038040713501147>
- Coburn, C. E., & Russell, J. L. (2008). District policy and teachers' social networks. *Educational Evaluation and Policy Analysis*, 30(3), 203–235. <https://doi.org/10.3102/0162373708321829>
- Common Core State Standards Initiative. (2020a). *Application of Common Core State Standards for English language learners*. <http://www.corestandards.org/assets/application-for-english-learners.pdf>
- Common Core State Standards Initiative. (2020b). *Application to students with disabilities*. <http://www.corestandards.org/wp-content/uploads/Application-to-Students-with-Disabilities-again-for-merge1.pdf>
- Common Core State Standards Initiative. (2020c). *Standards for mathematical practice*. <http://www.corestandards.org/Math/Practice/>
- Danielson, C. (2009). *Implementing the framework for teaching in enhancing professional practice*. ASCD.
- Desimone, L. (2002). How can comprehensive school reform models be successfully implemented? *Review of Educational Research*, 72(3), 433–479. <https://doi.org/10.3102/00346543072003433>
- Desimone, L. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181–199. <https://doi.org/10.3102/0013189X08331140>
- Desimone, L. M., & Garet, M. S. (2015). Best practices in teacher's professional development in the United States. *Psychology, Society, & Education*, 7(3), 252–263. <https://doi.org/10.25115/psy.v7i3.515>
- Desimone, L. M., Stornaiuolo, A., Flores, N., Pak, K., Edgerton, A., Nichols, T. P., Plummer, E. C., & Porter, A. (2019). Successes and challenges of the “new” college- and career-ready standards: Seven implementation trends. *Educational Researcher*, 48(3), 167–178. <https://doi.org/10.3102/0013189X19837239>

- Easton, J. Q., Johnson, E., & Sartain, L. (2017). *The predictive power of ninth-grade GPA*. The University of Chicago Consortium on School Research. <https://consortium.uchicago.edu/publications/predictive-power-ninth-grade-gpa>
- Egerton, A. K. (2020). Learning from standards deviations: Three dimensions for building education policies that last. *American Educational Research Journal*, 57(4), 1525–1566. <https://doi.org/10.3102/0002831219876566>
- Egerton, A. K., & Desimone, L. (2018). Teacher implementation of college- and career-ready standards: Links among policy, instruction, challenges, and resources. *AERA Open*, 4(4), 1–22. <https://doi.org/10.1177/2332858418806863>
- Fennema, E., Carpenter, T. P., Franke, M. L., Levi, L., Jacobs, V. R., & Empson, S. B. (1996). A longitudinal study of learning to use children's thinking in mathematics instruction. *Journal for Research in Mathematics Education*, 27(4), 403–434. <https://doi.org/10.2307/749875>
- Gamoran, A., & Hannigan, E. C. (2000). Algebra for everyone? Benefits of college-preparatory mathematics for students with diverse abilities in early secondary school. *Educational Evaluation and Policy Analysis*, 22(3), 241–254. <https://doi.org/10.3102/01623737022003241>
- Gao, N., & Lafortune, J. (2019). *Common Core State Standards in California: Evaluating local implementation and student outcomes*. Public Policy Institute of California. <https://www.ppic.org/publication/common-core-state-standards-in-california-evaluating-local-implementation-and-student-outcomes/>
- Gwynne, J. A., & Cowhy, J. R. (2017). *Getting ready for the Common Core State Standards: Experiences of CPS teachers and administrators preparing for the new standards*. University of Chicago Consortium on School Research.
- Haager, D., & Vaughn, S. (2013). The common core state standards and reading: Interpretations and implications for elementary students with learning disabilities. *Learning Disabilities Research & Practice*, 28(1), 5–16. <https://doi.org/10.1111/ldrp.12000>
- Hamilton, L. S., Kaufman, J. H., Stecher, B. M., Naftel, S., Robbins, M., Thompson, L. R., Garber, C., Faxon-Mills, S., & Opfer, D. O. (2016). *What supports do teachers need to help students meet Common Core State Standards for Mathematics? Findings from the American Teacher and American School Leader Panels*. RAND Corporation, 2016. https://www.rand.org/pubs/research_reports/RR1404-1.html
- Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
- Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1–28). Lawrence Erlbaum.
- Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology*, 45(3), 850. <https://doi.org/10.1037/a0014939>
- Kane, T. J., Owens, A. M., Marinell, W. H., Thal, D. R. C., & Staiger, D. O. (2016). *Teaching higher: Educators' perspectives on Common Core implementation*. Center for Education Policy Research, Harvard University. <https://cepr.harvard.edu/teaching-higher>
- Kazemi, E. (1998). Discourse that promotes conceptual understanding. *Teaching Children Mathematics*, 4(7), 410–414. <https://doi.org/10.5951/TCM.4.7.0410>
- Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, 86(4), 948–980. <https://doi.org/10.3102/0034654315626800>
- Klugman, J., Gordon, M. F., Sebring, P. B., & Sporte, S. E. (2015). A first look at the 5Essentials in Illinois schools (Research Report). *University of Chicago Consortium on Chicago School Research*. <https://consortium.uchicago.edu/publications/first-look-5essentials-illinois-schools>
- Kober, N., McIntosh, S., & Rentner, D. S. (2013). Year 3 of implementing the Common Core State Standards: Professional development for teachers and principals. *Center on Education Policy*. <https://files.eric.ed.gov/fulltext/ED555411.pdf>
- Lent, R. W., Brown, S. D., & Larkin, K. C. (1984). Relation of self-efficacy expectations to academic achievement and persistence. *Journal of Counseling Psychology*, 31(3), 356–362. <https://doi.org/10.1037/0022-0167.31.3.356>
- Loveless, T. (2014). *A progress report on the Common Core*. Brookings Institution. <https://www.brookings.edu/research/a-progress-report-on-the-common-core/>
- Loveless, T. (2016). *How well are American students learning* (The Brown Center Report on Education). <https://www.brookings.edu/wp-content/uploads/2016/03/Brown-Center-Report-2016.pdf>
- McLaughlin, M., Glaab, L., & Carrasco, I. H. (2014). *Implementing Common Core State Standards in California: A report from the field*. Policy Analysis for California Education. <https://files.eric.ed.gov/fulltext/ED564292.pdf>
- National Council of Teachers of Mathematics. (2018). *Catalyzing change in high school mathematics*.
- Newmann, F. M., Bryk, A. S., & Nagaoka, J. K. (2001). *Authentic intellectual work and standardized tests: Conflict or coexistence?* Consortium on Chicago School Research. <https://consortium.uchicago.edu/publications/authentic-intellectual-work-and-standardized-tests-conflict-or-coexistence>
- Niemi, D. (1996). Assessing conceptual understanding in mathematics: Representations, problem solutions, justifications, and explanations. *Journal of Educational Research*, 89(6), 351–363. <https://doi.org/10.1080/00220671.1996.9941339>
- Nomi, T., & Allensworth, E. (2013). Sorting and supporting: Why double-dose algebra led to better test scores but more course failures. *American Educational Research Journal*, 50(4), 756–788. <https://doi.org/10.3102/0002831212469997>
- Opfer, V. D., Kaufman, J. H., & Thompson, L. E. (2016). *Implementation of K–12 State Standards for mathematics and English language arts and literacy: Findings from the American Teacher Panel*. RAND Corporation. https://www.rand.org/content/dam/rand/pubs/research_reports/RR1500/RR1529-1/RAND_RR1529-1.pdf
- Opfer, V. D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, 81(3), 376–407. <https://doi.org/10.3102/0034654311413609>
- Polikoff, M. S. (2017). Is Common Core “working”? And where does Common Core research go from here? *AERA Open*, 3(1). <https://doi.org/10.1177/2332858417691749>
- Polikoff, M. S., & Porter, A. C. (2014). Instructional alignment as a measure of teaching quality. *Educational Evaluation and*

- Policy Analysis*, 36(4), 399–416. <https://doi.org/10.3102/0162373714531851>
- Powell, S. R., Fuchs, L. S., & Fuchs, D. (2013). Reaching the mountaintop: Addressing the common core standards in mathematics for students with mathematics difficulties. *Learning Disabilities Research & Practice*, 28(1), 38–48. <https://doi.org/10.1111/ldrp.12001>
- Prawat, R. S. (1989). Teaching for understanding: Three key attributes. *Teaching and Teacher Education*, 5(4), 315–328. [https://doi.org/10.1016/0742-051X\(89\)90029-2](https://doi.org/10.1016/0742-051X(89)90029-2)
- Rentner, D. S., & Kober, N. (2014). *Common Core State Standards in 2014: Districts' perceptions, progress, and challenges*. Center on Education Policy. <https://files.eric.ed.gov/fulltext/ED555389.pdf>
- Richland, L., Stigler, J. W., & Holyoak, K. J. (2012). Teaching the conceptual structure of mathematics. *Educational Psychologist*, 47(3), 189–203. <https://doi.org/10.1080/00461520.2012.667065>
- Rittle-Johnson, B., & Schneider, M. (2015). Developing conceptual and procedural knowledge of mathematics. In R. C. Kadosh, & A. Dowker (Eds.), *The Oxford handbook of numerical cognition* (pp. 1118–1134). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199642342.013.014>
- Rittle-Johnson, B., Sielger, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, 93(2), 346–362. <https://doi.org/10.1037/0022-0663.93.2.346>
- Saunders, A. F., Bethune, K. S., Spooner, F., & Browder, D. (2013). Solving the Common Core equation: Teaching mathematics CCSS to students with moderate and severe disabilities. *Teaching Exceptional Children*, 45(3), 24–33. <https://doi.org/10.1177/004005991304500303>
- Schoenfeld, A. H., & Teaching for Robust Understanding Project. (2016). *An introduction to the teaching for robust understanding (TRU) framework*. Graduate School of Education. <http://map.mathshell.org/trumath.php>
- Scholastic & Bill & Melinda Gates Foundation. (2014). *Teachers' views on the Common Core State Standards one year later*. <http://www.scholastic.com/primarysources/PrimarySources-2014update.pdf>
- Schweig, J. D., Kaufman, J. H., & Opfer, D. V. (2020). Day by day: Investigating variation in elementary mathematics instruction that supports the Common Core. *Educational Researcher*, 49(3), 176–187. <https://doi.org/10.3102/0013189X20909812>
- Sherhoff, D., Csikszentmihalyi, M., Shneider, B., & Sherhoff, E. S. (2003). Student engagement in high school classrooms from the perspective of flow theory. *School Psychology Quarterly*, 18(2), 158–176. <https://doi.org/10.1521/scpq.18.2.158.21860>
- Simzar, R., Domina, T., & Tran, C. (2016). Eighth-grade algebra course placement and student motivation for mathematics. *AERA Open*, 2(1). <https://doi.org/10.2332858415625227>
- Song, M., Yang, R., & Garet, M. (2019). *Effects of states' implementation of college- and career-ready standards on student achievement*. Center on Standards, Alignment, Instruction, and Learning.
- Toch, T. (1991). *In the name of excellence: The struggle to reform the nation's schools, why it's failing, and what should be done*. Oxford University Press.
- Tyack, D., & Cuban, L. (1995). *Tinkering toward utopia: A century of public school reform*. Harvard University Press.
- Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performance. *Education Policy Analysis Archives*, 10(12), 1–30. <https://doi.org/10.14507/epaa.v10n12.2002>
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. MESA Press.
- Xu, Z., & Cepa, K. (2018). Getting college-ready during state transition toward the Common Core State Standards. *Teachers College Record*, 120(6). <https://eric.ed.gov/?id=ED557605>

Authors

ELAINE ALLENSWORTH is the Lewis-Sebring Director of the University of Chicago Consortium on Chicago School Research. Her research examines factors influencing students' educational attainment, school leadership, and school improvement.

SARAH CASHDOLLAR is an Institute of Education Sciences predoctoral fellow in the Department of Comparative Human Development at the University of Chicago. She studies the impacts of Career and Technical Education and standards-based education reform on adolescent educational experiences and outcomes.

JULIA GWYNNE was a managing director at the University of Chicago Consortium on School Research when conducting this research. Currently she is a consultant at the Consortium. Her research interests include studies of educational policy and charter schools.

Appendix A

Student and Teacher Survey Items

Participation in CCSS-M-Related Professional Learning (2015 Teacher Survey):

How often did you receive formal training or professional development on CCSS so far this year?

(1-Never, 2-Once this year, 3-Once or twice a semester, 4-Monthly, 5-Weekly)

Which of the following topics have been addressed in your CCSS-M training and professional development? Check all that apply.

- Common Core Standards in Mathematics
- Curriculum materials and resources to teach the Common Core standards
- Teaching common standards to specific student groups (for example, students with disabilities or English Language Learners)
- Adapting classroom assessments to the Common Core standards
- New standardized assessments aligned with CCSS
- Research on best practices for implementation of the Common Core standards

Math Instructional Practices (2011-2017 Student Survey)

In your MATH class this year, how often do you do the following:

(1-Never, 2-Rarely, 3-Sometimes, 4-About half the time the class meets, 5-About every time the class meets)

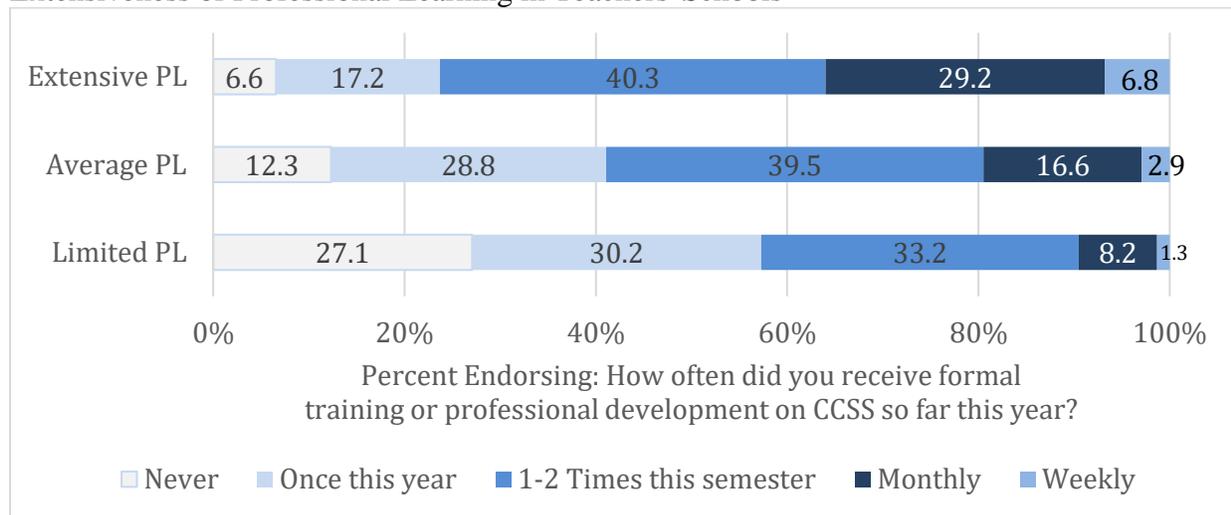
- Write a few sentences to explain how you solved a math problem.
- Explain how you solved a problem to the class.
- Write a math problem for other students to solve.
- Discuss possible solutions to problems with other students.
- Apply math to situations in life outside of school.
- Solve a problem with multiple steps that takes more than 20 minutes.
- Write a few sentences to explain how you solved a math problem.

Appendix B

Further Information on Schools in the Professional Learning Categories

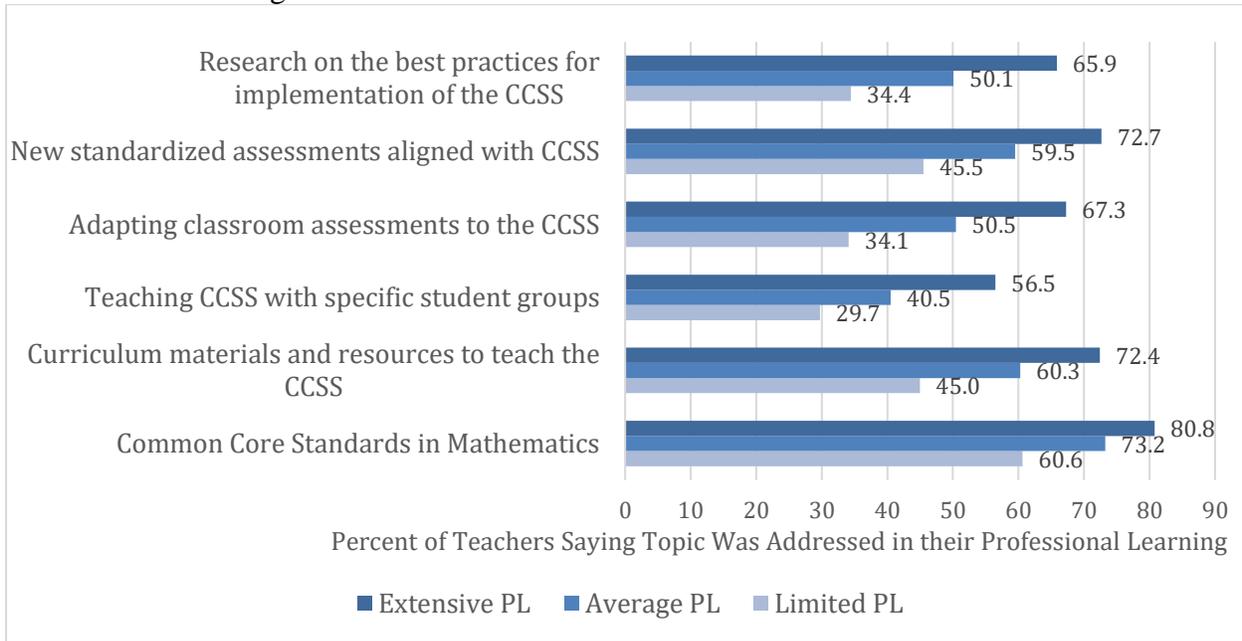
At schools with extensive professional learning around the standards, about 35 percent of teachers reported engaging in formal PL around the CCSS at least monthly, while only 25 percent reported engaging in PL once a year or less (see Figure B1). In contrast, at schools with limited PL, less than 10 percent of teachers engaged in PL around the standards at least once a month, while almost 60 percent did so once a year or less. About 70 percent of teachers at schools with extensive PL around the standards reported addressing topics that included research on best practices for implementation of the CCSS, new assessments aligned with CCSS, adapting classroom assessments to the CCSS, and curriculum materials and resources to teach the CCSS. At schools with limited PL, fewer than half of teachers said those topics were addressed in their professional learning.

Figure B1. Frequency of Professional Development around the CCSS-M by School the Overall Extensiveness of Professional Learning in Teachers' Schools



IMPROVEMENTS IN MATH INSTRUCTION AND ACHIEVEMENT WITH CCSS-M

Figure B2. Teacher Reports of CCSS Topics in their Professional Learning by Extensiveness of Professional Learning in their School



Appendix C

Results from Models with Different Specifications

Replication with Full Data

To ensure our conclusions were not affected by limiting the analytic sample to students with data on all outcomes we also ran the models inclusive of all students who had data on any outcome. For the analysis of the math practices survey measure, using the full data allows students at charter schools to be included in the analysis. For the analysis of NWEA scores, using the full data allows students at charter schools and students who did not take the spring survey to be included. For the analysis of grades and pass rates, using the full data allows students who left the district before the spring test or the spring survey to be included--analyzing their fall semester grades. The results are shown in Tables C.1-C.3 below.

The patterns are similar, with significant coefficients for all of the outcomes in post-implementation years among low- and average-achieving students in schools with extensive professional learning, and mostly insignificant coefficients among high achieving students. However, the coefficients are larger and more consistently significant for the measures of math practices and NWEA scores among students with low and average prior-year achievement in the models with all cases than in the models with the reduced sample. As observed with the reduced sample, there were fewer and less consistent significant differences between schools with average PL and schools with limited PL than between those with extensive PL and limited PL.

IMPROVEMENTS IN MATH INSTRUCTION AND ACHIEVEMENT WITH CCSS-M

Table C.1. Results for Low Achieving Students with Full Data

	Coefficient	Math Practices		NWEA Score		GPA		Pass Math	
		Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
	Intercept	-0.103	0.023 ***	-0.210	0.020 ***	2.268	0.025 ***	0.981	0.007 ***
Year Fixed Effects (Relative to 2013)	2011 (Pre-CCSSM)	-0.179	0.039 ***			-0.223	0.052 ***	-0.065	0.016 ***
	2012 (Pre-CCSSM)	-0.099	0.034 **			-0.067	0.052	-0.020	0.013
	2013 (Pre-CCSSM)	-0.048	0.037	-0.096	0.023 ***	-0.010	0.032	-0.015	0.008 .
	2015 (Post)	0.073	0.036 *	0.129	0.019 ***	0.038	0.032	0.005	0.011
	2016 (Post)	0.059	0.038	0.123	0.024 ***	0.125	0.034 ***	0.025	0.009 **
	2017 (Post)	0.127	0.044 **	0.105	0.030 ***	0.152	0.037 ***	0.031	0.010 **
Schools with Average PL	2011xAveragePL (Pre)	0.036	0.048			0.127	0.062 *	0.035	0.019 .
	2012xAveragePL (Pre)	0.035	0.043			0.010	0.063	-0.002	0.017
	2013xAveragePL (Pre)	0.042	0.047	-0.026	0.032	0.000	0.044	0.000	0.012
	2015xAveragePL	0.039	0.044	-0.011	0.027	0.097	0.043 *	0.024	0.014 .
	2016xAveragePL	0.157	0.049 **	0.036	0.032	0.078	0.048	0.023	0.013 .
	2017xAveragePL	0.075	0.057	0.044	0.039	0.077	0.050	0.022	0.014
Schools with Extensive PL	2011xHighPL (Pre)	0.025	0.048			0.069	0.062	0.012	0.020
	2012xHighPL (Pre)	-0.018	0.043			0.018	0.061	-0.008	0.017
	2013xHighPL (Pre)	0.045	0.045	-0.046	0.030	-0.038	0.044	-0.013	0.012
	2015xHighPL	0.148	0.054 **	-0.003	0.026	0.126	0.043 **	0.025	0.013 .
	2016xHighPL	0.152	0.046 *	0.094	0.033 **	0.171	0.044 ***	0.039	0.013 **
	2017xHighPL	0.114	0.056 *	0.109	0.040 **	0.136	0.049 **	0.029	0.013 *
Covariates	Grade 7	-0.047	0.010 ***	0.238	0.010 ***	0.007	0.016	0.004	0.005
	Grade 8	-0.040	0.011 ***	0.533	0.010 ***	0.222	0.015 ***	0.043	0.004 ***
	Algebra	0.052	0.063	0.250	0.056 ***	-0.062	0.067	-0.029	0.010 **
	Other math	0.808	0.052 ***	0.915	0.031 ***	3.753	0.042 ***	0.957	0.012 ***
	Multiple math	-0.012	0.053	0.071	0.059	0.225	0.106 *	0.002	0.012
	White	-0.060	0.026 *	0.065	0.018 ***	0.231	0.019 ***	0.031	0.006 ***
	Asian	0.003	0.033	0.136	0.025 ***	0.365	0.032 ***	0.048	0.007 ***
	Hispanic	-0.123	0.014 ***	0.038	0.010 ***	0.105	0.014 ***	0.018	0.004 ***
	Other	-0.069	0.035 *	0.093	0.025 ***	0.126	0.032 ***	0.012	0.010
	Male	-0.044	0.008 ***	0.042	0.005 ***	-0.266	0.009 ***	-0.053	0.003 ***
	Poverty	0.005	0.006	-0.012	0.004 **	-0.010	0.005 .	0.000	0.002
	Social Status	-0.008	0.007	0.002	0.005	-0.002	0.006	0.000	0.002
	Prior math score	0.128	0.008 ***	0.701	0.011 ***	0.176	0.012 ***	0.032	0.003 ***

Note: .<.10, *<.05, **<.01, ***<.001. Low achievement is defined as at least 0.5 s.d. below the district mean in the prior year on the state assessment. Coefficients come from models that also include school fixed-effects. Standard errors were clustered by school. Math practices and NWEA scores were standardized using means and standard deviations from 2014. Teacher professional learning (PL) around the CCSS-M was measured in the 2014-15 school year, which was the same year all schools were to begin implementing the standards.

IMPROVEMENTS IN MATH INSTRUCTION AND ACHIEVEMENT WITH CCSS-M

Table C.2. Results for Average Achieving Students with Full Data

	Coefficient	Math Practices		NWEA Score		GPA		Pass Math	
		Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
	Intercept	-0.124	0.018 ***	-0.189	0.011 ***	2.666	0.020 ***	0.995	0.002 ***
Year Fixed Effects (Relative to 2013)	2011 (Pre-CCSSM)	-0.109	0.037 **	-0.079	0.019 ***	-0.246	0.062 ***	-0.036	0.009 ***
	2012 (Pre-CCSSM)	-0.074	0.033 *	0.012	0.014	-0.057	0.042	-0.004	0.005
	2013 (Pre-CCSSM)	0.030	0.030	0.021	0.020	-0.030	0.027	-0.010	0.003 **
	2015 (Post)	0.064	0.031 *	-0.016	0.026	-0.021	0.033	-0.006	0.004 .
	2016 (Post)	0.062	0.033 .			0.051	0.036	0.003	0.004
	2017 (Post)	0.067	0.037 .			0.025	0.039	0.003	0.005
Schools with Average PL	2011xAveragePL (Pre)	0.012	0.048			0.153	0.071	0.015	0.011
	2012xAveragePL (Pre)	0.039	0.042			-0.028	0.053	-0.017	0.008 *
	2013xAveragePL (Pre)	0.020	0.039	0.010	0.027	0.028	0.036	0.005	0.005
	2015xAveragePL	0.042	0.039	0.043	0.019 *	0.032	0.046	0.004	0.005
	2016xAveragePL	0.108	0.043 *	0.059	0.026 *	0.026	0.047	0.009	0.005 .
	2017xAveragePL	0.100	0.047 *	0.041	0.032	0.073	0.051	0.009	0.006
Schools with Extensive PL	2011xHighPL (Pre)	-0.056	0.045			0.101	0.070	0.009	0.011
	2012xHighPL (Pre)	-0.024	0.042			0.003	0.051	-0.017	0.007 *
	2013xHighPL (Pre)	-0.027	0.036	-0.018	0.026	0.013	0.039	0.002	0.005
	2015xHighPL	0.088	0.044 *	0.043	0.021 *	0.079	0.042 .	0.012	0.005 *
	2016xHighPL	0.103	0.041 *	0.094	0.028 ***	0.149	0.047 **	0.013	0.006 *
	2017xHighPL	0.124	0.048 **	0.103	0.034 **	0.157	0.049 **	0.015	0.006 *
Covariates	Grade 7	-0.036	0.010 **	0.305	0.010 ***	-0.048	0.018 **	-0.010	0.002 ***
	Grade 8	-0.048	0.011 ***	0.541	0.010 ***	0.072	0.017 ***	0.009	0.002 ***
	Algebra	0.027	0.018	0.217	0.018 ***	-0.065	0.028 *	-0.003	0.003
	Other math	-0.460	0.202 *	0.885	0.163 ***	0.404	0.343	0.059	0.017 ***
	Multiple math	0.049	0.026 .	0.124	0.024 ***	0.180	0.041 ***	-0.006	0.004
	White	-0.140	0.018 ***	0.098	0.011 ***	0.257	0.019 ***	0.011	0.003 ***
	Asian	-0.081	0.027 **	0.199	0.016 ***	0.508	0.031 ***	0.024	0.005 ***
	Hispanic	-0.148	0.012 ***	0.037	0.008 ***	0.127	0.014 ***	0.006	0.002 **
	Other	-0.089	0.026 ***	0.038	0.016 *	0.100	0.027 ***	-0.001	0.005
	Male	-0.033	0.007 ***	0.033	0.003 ***	-0.372	0.009 ***	-0.034	0.002 ***
	Poverty	0.012	0.005 *	-0.004	0.003	-0.010	0.005 *	-0.002	0.001 **
	Social Status	0.001	0.005	0.003	0.003	-0.009	0.005 .	-0.001	0.001
	Prior math score	0.114	0.010 ***	0.786	0.006 ***	0.730	0.011 ***	0.048	0.003 ***

Note: .<.10, *<.05, **<.01, ***<.001. Average achievement is defined as at within 0.5 s.d. of the district mean in the prior year on the state assessment. Coefficients come from models that also include school fixed-effects. Standard errors were clustered by school. Math practices and NWEA scores were standardized using means and standard deviations from 2014. Teacher professional learning (PL) around the CCSS-M was measured in the 2014-15 school year, which was the same year all schools were to begin implementing the standards.

IMPROVEMENTS IN MATH INSTRUCTION AND ACHIEVEMENT WITH CCSS-M

Table C.3. Results for High Achieving Students with Full Data

	Coefficient	Math Practices		NWEA Score		GPA		Pass Math	
		Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
	Intercept	-0.092	0.018 ***	-0.084	0.012 ***	2.961	0.016 ***	0.991	0.001 ***
Year Fixed Effects (Relative to 2013)	2011 (Pre-CCSSM)	-0.029	0.034			-0.164	0.032 ***	-0.006	0.002 **
	2012 (Pre-CCSSM)	-0.019	0.033			-0.108	0.032 ***	-0.003	0.002
	2013 (Pre-CCSSM)	0.098	0.031 **	-0.082	0.020 ***	-0.047	0.020 *	-0.001	0.001
	2015 (Post)	0.070	0.026 **	0.081	0.012 ***	-0.013	0.021	0.001	0.001
	2016 (Post)	0.104	0.027 ***	0.143	0.023 ***	0.080	0.026 **	0.001	0.002
	2017 (Post)	0.048	0.034	0.191	0.025 ***	0.074	0.030 *	0.004	0.001
Schools with Average PL	2011xAveragePL (Pre)	0.019	0.048			0.071	0.041 .	0.002	0.003
	2012xAveragePL (Pre)	0.006	0.045			0.061	0.039	0.000	0.003
	2013xAveragePL (Pre)	-0.029	0.040	-0.008	0.025	0.032	0.030	0.000	0.002
	2015xAveragePL	0.019	0.035	0.021	0.018	0.018	0.029	0.000	0.002
	2016xAveragePL	0.086	0.038 *	0.037	0.028	-0.019	0.034	0.003	0.002
	2017xAveragePL	0.097	0.046 *	0.016	0.033	0.013	0.038	0.000	0.002
Schools with Extensive PL	2011xHighPL (Pre)	-0.053	0.044			0.059	0.041	0.002	0.003
	2012xHighPL (Pre)	-0.028	0.043			0.083	0.040 *	0.000	0.003
	2013xHighPL (Pre)	-0.089	0.040 *	0.018	0.026	0.046	0.030	0.000	0.002
	2015xHighPL	0.071	0.035 *	0.021	0.021	0.036	0.030	0.001	0.002
	2016xHighPL	0.018	0.037	0.042	0.030	0.049	0.035	0.002	0.002
	2017xHighPL	0.074	0.046	0.064	0.035 .	0.086	0.038 *	0.000	0.002
Covariates	Grade 7	-0.029	0.010 **	0.390	0.011 ***	0.010	0.015	-0.002	0.001 **
	Grade 8	-0.086	0.014 ***	0.645	0.012 ***	-0.037	0.018 *	0.000	0.001
	Algebra	-0.005	0.018	0.177	0.014 ***	-0.057	0.019 **	0.001	0.001
	Other math	-0.043	0.077	0.342	0.064 ***	0.360	0.129 **	0.026	0.018
	Multiple math	0.065	0.023 **	0.122	0.019 ***	0.108	0.023 ***	-0.001	0.001
	White	-0.074	0.016 ***	0.138	0.011 ***	0.225	0.018 ***	0.004	0.001 **
	Asian	-0.040	0.017 *	0.202	0.016 ***	0.330	0.022 ***	0.006	0.002 ***
	Hispanic	-0.081	0.014 ***	0.038	0.011 ***	0.117	0.016 ***	0.003	0.001 *
	Other	-0.059	0.020 *	0.132	0.014 ***	0.160	0.023 ***	0.003	0.002 *
	Male	-0.026	0.007 ***	0.052	0.004 ***	-0.287	0.007 ***	-0.006	0.001 ***
	Poverty	0.005	0.005	-0.014	0.003 ***	-0.023	0.004 ***	0.000	0.000
	Social Status	0.002	0.005	0.011	0.004 **	0.006	0.005	0.000	0.001
Prior math score	0.040	0.006 ***	0.590	0.007 ***	0.429	0.009 ***	0.005	0.001 ***	

Note: .<.10, *<.05, **<.01, ***<.001. High achievement is defined as at least 0.5 s.d. above the district mean in the prior year on the state assessment. Coefficients come from models that also include school fixed-effects. Standard errors were clustered by school. Math practices and NWEA scores were standardized using means and standard deviations from 2014. Teacher professional learning (PL) around the CCSS-M was measured in the 2014-15 school year, which was the same year all schools were to begin implementing the standards.

Professional Learning as a Continuous Variable

The models were also run with professional learning as a continuous variable, instead of separating schools into three groups based on the extensiveness of their professional learning. Table C4 shows the results with all students combined, while Table C5 shows the results for models separated by prior student achievement. The coefficients that emerge as significant for the continuous version are similar to those seen for schools with extensive professional learning, although the coefficient on the continuous PL variable predicting NWEA scores is not significant until the 2017 year. In the subgroup analyses shown in Table C6, the patterns of which of the continuous-PL coefficients are significant is also similar to those observed among schools with extensive professional development in the prior tables, with many reaching a level of statistical significance for low- and average-achieving students, but not for high-achieving students.

Table C4. Professional Learning as a Continuous Variable Predicting Changes in Student Outcomes: All Students

Coefficient	Math Practices		NWEA Score		GPA		Pass Math	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Intercept	-0.083	0.011 ***	-0.151	0.008 ***	2.711	0.011 ***	0.986	0.002 ***
2011 (Pre-CCSSM)	-0.121	0.016 ***			-0.133	0.017 ***	-0.024	0.003 ***
2012 (Pre-CCSSM)	-0.072	0.014 ***			-0.058	0.016 ***	-0.014	0.003 ***
2013 (Pre-CCSSM)	0.017	0.013	-0.092	0.010 ***	-0.019	0.013	-0.008	0.002 ***
2015 (Post)	0.121	0.015 ***	0.090	0.008 ***	0.059	0.013 ***	0.010	0.002 ***
2016 (Post)	0.157	0.014 ***	0.158	0.011 ***	0.147	0.015 ***	0.022	0.002 ***
2017 (Post)	0.162	0.017 ***	0.180	0.013 ***	0.154	0.016 ***	0.020	0.003 ***
2011 x PL (Pre)	-0.035	0.019 .			0.010	0.017	0.001	0.004
2012 x PL (Pre)	-0.026	0.019			0.008	0.016	-0.003	0.003
2013 x PL (Pre)	-0.020	0.015	-0.020	0.010 *	0.000	0.013	-0.002	0.003
2015 x PL	0.039	0.017 *	0.002	0.010	0.033	0.014 *	0.007	0.002 **
2016 x PL	0.029	0.014 *	0.016	0.015	0.056	0.016 ***	0.012	0.003 ***
2017 x PL	0.037	0.019 .	0.030	0.015 *	0.041	0.017 *	0.009	0.003 ***
Grade 7	-0.038	0.008 ***	0.319	0.010 ***	-0.013	0.014	-0.002	0.002
Grade 8	-0.048	0.010 ***	0.580	0.010 ***	0.106	0.014 ***	0.023	0.002 ***
Algebra	-0.022	0.015	0.222	0.014 ***	-0.124	0.019 ***	-0.020	0.002 ***
Other math	-0.127	0.097	0.553	0.069 ***	0.167	0.117	-0.002	0.017
Multiple math	0.056	0.021 **	0.122	0.022 ***	0.159	0.032 ***	-0.001	0.004
White	-0.114	0.011 ***	0.096	0.010 ***	0.230	0.013 ***	0.005	0.003 .
Asian	-0.078	0.015 ***	0.165	0.013 ***	0.369	0.021 ***	0.006	0.004 .
Hispanic	-0.133	0.009 ***	0.028	0.008 ***	0.111	0.010 ***	0.008	0.002 ***
Other	-0.087	0.016 ***	0.073	0.012 ***	0.127	0.018 ***	0.002	0.003
Male	-0.042	0.005 ***	0.041	0.003 ***	-0.308	0.006 ***	-0.032	0.001 ***
Poverty	0.007	0.004 *	-0.011	0.002 ***	-0.015	0.003 ***	-0.001	0.001
Social Status	0.002	0.004	0.003	0.003	-0.002	0.004	-0.001	0.001
Prior math score	0.096	0.003 ***	0.758	0.005 ***	0.558	0.006 ***	0.036	0.002 ***
R-square	0.052		0.714		0.379		0.078	

Note: .<.10, *<.05, **<.01, ***<.001. Coefficients come from models that also include school fixed-effects. Standard errors were clustered by school. Math practices and NWEA scores were standardized using means and standard deviations from 2014. Teacher professional learning (PL) around the CCSS-M was measured in the 2014-15 school year, which was the same year all schools were to begin implementing the standards. It was standardized around across school means prior to analysis.

Table C5. Professional Learning as a Continuous Variable Predicting Changes in Student Outcomes: DiD Coefficients for Subgroups of Students by Prior Achievement

	Coefficient	Math Practices			NWEA Score			GPA			Pass Math		
		Estimate	S.E.	p-val	Estimate	S.E.	p-val	Estimate	S.E.	p-val	Estimate	S.E.	p-val
Low-Achieving Students	2011 x PL (Pre)	-0.015	0.024	0.53				0.000	0.024	1.00	0.004	0.009	0.69
	2012 x PL (Pre)	-0.016	0.023	0.50				0.011	0.024	0.65	0.001	0.007	0.93
	2013 x PL (Pre)	0.007	0.022	0.76	-0.030	0.014	0.03	-0.011	0.019	0.57	-0.003	0.006	0.57
	2015 x PL	0.058	0.028	0.04	0.008	0.015	0.61	0.047	0.020	0.02	0.012	0.006	0.04
	2016 x PL	0.048	0.023	0.04	0.035	0.018	0.05	0.073	0.022	0.00	0.021	0.007	0.00
2017 x PL	0.032	0.029	0.28	0.050	0.021	0.02	0.055	0.023	0.02	0.016	0.007	0.03	
Average-Achieving Students	2011 x PL (Pre)	-0.048	0.025	0.05				0.010	0.023	0.68	0.002	0.004	0.63
	2012 x PL (Pre)	-0.029	0.023	0.20				-0.004	0.020	0.85	-0.006	0.003	0.04
	2013 x PL (Pre)	-0.026	0.018	0.14	-0.021	0.012	0.08	-0.001	0.018	0.97	-0.002	0.002	0.47
	2015 x PL	0.025	0.020	0.21	0.007	0.012	0.57	0.027	0.019	0.16	0.005	0.003	0.05
	2016 x PL	0.028	0.019	0.13	0.011	0.016	0.49	0.058	0.021	0.01	0.007	0.003	0.03
2017 x PL	0.039	0.022	0.08	0.044	0.018	0.01	0.043	0.023	0.06	0.006	0.003	0.08	
High-achieving students	2011 x PL (Pre)	-0.025	0.022	0.26				0.021	0.019	0.27	0.001	0.001	0.59
	2012 x PL (Pre)	-0.026	0.022	0.24				0.017	0.016	0.30	-0.001	0.001	0.51
	2013 x PL (Pre)	-0.031	0.019	0.11	-0.008	0.013	0.53	0.012	0.013	0.36	0.000	0.001	0.77
	2015 x PL	0.031	0.018	0.09	-0.008	0.011	0.48	0.007	0.013	0.57	0.000	0.001	0.90
	2016 x PL	0.007	0.017	0.67	0.000	0.015	0.98	0.017	0.016	0.29	0.001	0.001	0.32
2017 x PL	0.027	0.022	0.21	0.013	0.015	0.40	0.017	0.017	0.33	0.000	0.001	0.81	

Note: Coefficients come from models that also include school fixed-effects. Standard errors were clustered by school. Math practices and NWEA scores were standardized using means and standard deviations from 2014. Teacher professional learning (PL) around the CCSS-M was measured in the 2014-15 school year, which was the same year all schools were to begin implementing the standards. It was standardized around across school means (mean and s.d. of school means) prior to analysis.

PARCC Scores as an Outcome

In 2015, Illinois switched its state test to the PARCC, which was developed in response to the CCSS-M. Because there are no scores available prior to standards implementation, we cannot conduct DiD analyses to examine changes over time in PARCC scores relative to pre-implementation years. Instead, we examine whether schools differed in their average PARCC scores in 2015, and in their growth in PARCC scores from 2015 to 2016 and 2017, controlling for student covariates, based on their level of PL in 2015. We do not use school fixed effects, as these would be estimated only on post-implementation years for schools in each of the PL implementation groups, and would be confounded with membership in the implementation groups. Instead, we use hierarchical linear models, with students nested within schools (see formula below).

IMPROVEMENTS IN MATH INSTRUCTION AND ACHIEVEMENT WITH CCSS-M

The models estimate a random effect for each school (u_{0j}), which is the average difference in PARCC scores for each school, controlling for the variables in the models. β_{0j} captures the school average PARCC score in 2015, adjusted for student covariates. β_{1j} and β_{2j} capture the deviation for each school from their 2015 mean. At the school level, γ_{00} is the average PARCC score in 2015 for schools with low levels of PL in 2015, while γ_{01} and γ_{02} are the differences in PARCC scores in 2015 for schools with average and extensive PL, relative to those with limited PL, controlling for average school achievement in 2014 on the prior state test (the ISAT). Likewise, γ_{11} and γ_{12} are the differences in the deviation from 2015 for schools with average and extensive PL, relative to those with limited PL, controlling for average school achievement in 2014 on the ISAT, with the 2017 deviation modeled similarly.

Level-1 Model

$$\text{PARCC}_{ij} = \beta_{0j} + \beta_1*(D16)_{ij} + \beta_2*(D17)_{ij} + \beta_3*(CGR7)_{ij} + \beta_4*(CGR8)_{ij} + \beta_5*(CALGEBRA)_{ij} + \beta_6*(COTH)_{ij} + \beta_7*(CMULTI)_{ij} + \beta_8*(CDWHITE)_{ij} + \beta_9*(CDASIAN)_{ij} + \beta_{10}*(CDHISP)_{ij} + \beta_{11}*(CDOOTHER)_{ij} + \beta_{12}*(CDMALE)_{ij} + \beta_{13}*(CSCON)_{ij} + \beta_{14}*(CSSOC)_{ij} + \beta_{15j}*(PRIORM)_{ij} + r_{ij}$$

Level-2 Model

$$\begin{aligned}\beta_{0j} &= \gamma_{00} + \gamma_{01}*(\text{Ave PL})_j + \gamma_{02}*(\text{High PL})_j + \gamma_{03}*(2014\text{SchoolAch})_j + u_{0j} \\ \beta_1 &= \gamma_{10} + \gamma_{11}*(\text{Ave PL})_j + \gamma_{12}*(\text{High PL})_j + \gamma_{13}*(2014\text{SchoolAch})_j \\ \beta_2 &= \gamma_{20} + \gamma_{21}*(\text{Ave PL})_j + \gamma_{22}*(\text{High PL})_j + \gamma_{23}*(2014\text{SchoolAch})_j \\ \beta_3 &= \gamma_{30} \\ &\dots \\ \beta_{15} &= \gamma_{150}\end{aligned}$$

As shown in Table C6, schools in the three PL groups did not differ in their PARCC scores in 2015, controlling for the covariates; neither the average nor extensive PL coefficients are significantly different from zero on the intercept (2015). In 2016, schools with limited PL did not show significant improvement in PARCC scores relative to their scores in 2015 (insignificant coefficient of 0.014 s.d.). Schools with average and extensive PL showed significantly larger improvements in that year than schools with limited PL (coefficients of 0.025 and 0.038,

IMPROVEMENTS IN MATH INSTRUCTION AND ACHIEVEMENT WITH CCSS-M

respectively; total change of 0.039 and 0.052, respectively). In 2017, schools with extensive professional learning in 2015 continued to show significantly higher gains in PARCC scores than schools with limited PL. However, in 2017 schools with average PL no longer showed higher PARCC scores than schools with limited PL, which differs from NWEA scores. It is not clear whether the difference results because of the outcome being measured (PARCC vs NWEA), the lack of pre-implementation data, different methodology, or different base year. But the general pattern is similar—significant differences did not appear until 2016, when they are seen among both schools with average and extensive PL, and they remained significant among schools with extensive PL in 2017, at about the same level as in 2016.

Table C6. Models Predicting PARCC Scores

Fixed Effects	Coeff	s.e.
Intercept (2015, Limited PL)	-0.014	0.015
Average PL (relative to Limited PL)	0.012	0.018
Extensive PL (relative to Limited PL)	0.000	0.018
2014 mean achievement	0.145	0.014 ***
Year 2016 (relative to 2015)	0.014	0.010
Average PL (relative to Limited PL)	0.025	0.012 *
Extensive PL (relative to Limited PL)	0.038	0.012 **
2014 mean achievement	-0.006	0.010
Year 2017 (relative to 2015)	-0.014	0.009
Average PL (relative to Limited PL)	0.000	0.012
Extensive PL (relative to Limited PL)	0.035	0.012 **
2014 mean achievement	0.050	0.010 **
Grade 7	0.061	0.004 ***
Grade 8	-0.112	0.005 ***
Algebra	0.414	0.008 ***
Other math	0.652	0.064 ***
Multiple math	0.193	0.012 ***
White	0.168	0.009 ***
Asian	0.256	0.012 ***
Hispanic	0.074	0.007 ***
Other	0.104	0.017 ***
Male	-0.049	0.004 ***
Poverty	-0.008	0.003 ***
Social Status	-0.003	0.003
Prior math score	0.752	0.002 ***
Random Components	Variance	in SD
Intercept, u ₀	0.014	0.119 ***
Residual, r	0.352	0.594 ***

* $p < .05$, ** $p < .01$, *** $p < .001$. PARCC scores are standardized based on the 2015 mean and standard deviation across students. Student-level prior math score is the prior state test (ISAT) for students in 2015 (2014 ISAT scores) and the PARCC in later years, each standardized within year and grade. School-level mean achievement is based on the 2014 ISAT. Analysis was conducted through hierarchical linear models with students nested within schools.

Including Teacher Fixed Effects

An additional specification adds teacher fixed effects to the models. Adding teacher fixed effects could address any concerns that there were changes in the teacher workforce in post-implementation years that account for changing instruction in schools with more professional

IMPROVEMENTS IN MATH INSTRUCTION AND ACHIEVEMENT WITH CCSS-M

development around the standards. However, it also can confound teacher and school effects, as some teachers only have observations in a small number of years, and many do not have observations in both pre-implementation and post-implementation years. To the extent that the middle grade teachers in a school in post-implementation years were different than those in the school in pre-policy years, it becomes difficult to attribute changes in practices and outcomes in the school to teachers versus school factors.

Table C7 shows that 40% of teachers in the sample only are observed in one of the seven years. This occurs not only because teachers may leave or be hired in different schools, but also because teachers move in and out of teaching in grades 6-8 (most CPS elementary schools span grades K-8). While this is a large percentage of teachers in the sample, it is only eight percent of the observations; the more years a teacher is present the more observations they can have in the sample, and long-term teachers tend to have more students in a given year than teachers who are observed just once. Only 36 percent of teachers taught students in the sample in both pre-implementation years (2014 and earlier) and post-implementation years (2015 and later). At the same time, 73 percent of the observations were attached to teachers that had observations in at least one pre-implementation and one post-implementation year. The further the observation was from 2014/2015, the less likely the teacher was observed in both years. The coefficients and standard errors should be interpreted cautiously not only because of the issues described above, but also because it is difficult to interpret the coefficients. The DiD coefficients are framed in reference to the 2014 year, and more than a quarter of all teachers did not teach in 2014, or any pre-policy year, making interpretation of coefficients and standard errors ambiguous.

One additional difference from the model without teacher fixed effects (Table C4) is that we only could include one teacher record per observation. In the prior analysis, we averaged

together grades and pass rates for students who took multiple math courses in a year, but it would not make sense to average teacher ID numbers. For this analysis, we deleted observations that were taken in math classes other than standard mathematics or algebra, so that we would use the record from a student's primary math class for the analysis. We also deleted any observations in which multiple teacher IDs were assigned to the same class (e.g., classes that were co-taught). These two restrictions only reduced the number of cases by 2.5 percent of the total.

Results from the models are shown in Table C8; they can be compared to those in Table C4 which is the same specification but without teacher fixed effects. The sizes of the DiD coefficients are similar in the two tables, although fewer of the coefficients are significant. The GPA and passing coefficients representing 2016 and 2017 remain significant in this specification, as well as the math practices measure in 2017. Strong similarity in the size of the coefficients in Table C4 provides evidence that the findings on grades and practices were not simply a result of changes in the teacher workforce. Changes in the significance levels in the model with teacher fixed effects could result from the lack of within-teacher variation in each school year, which is the major component of the DiD estimates.

Consideration of Student Fixed Effects. Readers may be interested in model specifications that include student fixed-effects, which would provide estimates of how individual students' outcomes changed relative to themselves in other years that they were in grades 6, 7 or 8. Because we only study the middle grades, even in the complete data set that includes all observations, only 27 percent of observations are based on students who have data in both pre-implementation and post-implementation years (see Table C9). The observations that span pre- and post-implementation periods are more likely to occur in the years immediately before and after implementation (2014 and 2015), so that estimates for 2016 and 2017 are based

IMPROVEMENTS IN MATH INSTRUCTION AND ACHIEVEMENT WITH CCSS-M

on small numbers of students. Only the small percentage of students held back in grade have observations for both 2014 and 2017. Because students must have at least two observations to be included in the analysis, the seventh-grade year becomes over-weighted relative to other years. For these reasons, adding student fixed-effects to the models makes the year and DiD coefficients uninterpretable.

Table C7. Number of years each teacher is observed in the sample

Number of years teacher is observed in the sample	Number of Teachers	Percent of Teachers	Number of Student Observations	Percent of Observations
1 year	1301	40%	24,481	8%
2 years	634	19%	32,265	11%
3 years	372	11%	32,227	11%
4 years	305	9%	38,962	13%
5 years	226	7%	41,488	14%
6 years	178	5%	48,360	16%
7 years	237	7%	86,865	29%

Table C8. Models with Teacher Fixed Effects and Continuous Measure of Professional Learning

	Math Practices			NWEA Score			GPA			Pass Math		
	Estimate	S.E.		Estimate	S.E.		Estimate	S.E.		Estimate	S.E.	
Intercept	-0.002	0.01		-0.026	0.01	***	2.532	0.01	***	0.955	0.002	***
2011 (Pre-CCSSM)	-0.099	0.02	***				-0.094	0.02	***	-0.025	0.004	***
2012 (Pre-CCSSM)	-0.061	0.02	***				-0.038	0.01	*	-0.014	0.003	***
2013 (Pre-CCSSM)	0.022	0.01		-0.095	0.01	***	-0.015	0.01		-0.009	0.002	***
2015 (Post)	0.113	0.01	***	0.092	0.01	***	0.043	0.01	**	0.006	0.002	**
2016 (Post)	0.143	0.02	***	0.169	0.01	***	0.126	0.01	***	0.017	0.003	***
2017 (Post)	0.148	0.02	***	0.194	0.02	***	0.136	0.02	***	0.017	0.003	***
2011 x PL (Pre)	-0.044	0.02	.				0.000	0.02		-0.005	0.004	
2012 x PL (Pre)	-0.032	0.02					-0.001	0.02		-0.006	0.003	.
2013 x PL (Pre)	-0.017	0.02		-0.030	0.01	**	-0.008	0.01		-0.002	0.003	
2015 x PL	0.028	0.02		0.000	0.01		0.022	0.02		0.004	0.003	
2016 x PL	0.024	0.02		0.017	0.02		0.046	0.02	**	0.008	0.003	**
2017 x PL	0.040	0.02	*	0.024	0.02		0.061	0.02	**	0.009	0.004	*
Grade 7	-0.026	0.01	**	0.339	0.01	***	-0.010	0.01		0.001	0.002	
Grade 8	-0.040	0.01	**	0.584	0.01	***	0.124	0.01	***	0.026	0.003	***
Algebra	-0.022	0.01		0.221	0.01	***	-0.134	0.02	***	-0.021	0.002	***
Multiple math	0.060	0.02	**	0.139	0.02	***	0.160	0.03	***	-0.002	0.004	
White	-0.115	0.01	***	0.097	0.01	***	0.207	0.01	***	0.003	0.003	
Asian	-0.076	0.01	***	0.156	0.01	***	0.356	0.02	***	0.005	0.004	
Hispanic	-0.127	0.01	***	0.031	0.01	***	0.103	0.01	***	0.007	0.002	**
Other	-0.078	0.02	***	0.075	0.01	***	0.108	0.02	***	0.000	0.003	
Male	-0.041	0.01	***	0.051	0	***	-0.323	0.01	***	-0.034	0.002	***
Poverty	0.008	0	*	-0.010	0	***	-0.017	0	***	-0.001	7E-04	
Social Status	0.004	0		0.003	0		-0.004	0		-0.001	8E-04	
Prior math score	0.088	0	***	0.729	0	***	0.617	0.01	***	0.043	0.002	***
R-square	0.089			0.746			0.453			0.127		

*p<.05, **p<.01, ***p<.001. Coefficients come from models that include school and teacher fixed-effects. Standard errors were clustered by school. Math practices and NWEA scores were standardized using means and standard deviations from 2014. Teacher professional learning (PL) around the CCSS-M was measured in the 2014-15 school year, which was the same year all schools were to begin implementing the standards. It was standardized around across school means (mean and s.d. of school means) prior to analysis.

Table C9. Number of years each student is observed in the full sample

Number of years student is observed in the sample	Number of Students	Percent of Students	Percent of total observations		Percent of total observations spanning pre- and post-implementation	
			Number of Observations	Percent of Observations	spanning pre- and post-implementation	spanning pre- and post-implementation
1 time	75,416	13.8%	75,416	5.5%	0%	0%
2 times	135,081	24.7%	270,162	19.9%	8%	2%
3 times	329,775	60.4%	989,325	72.8%	40%	24%
4 times	6,137	1.1%	24,548	1.8%	79%	1%
5 times	15	0.0%	75	0.0%	100%	0%

