Teacher Development and Mathematics Performance Across Ethnic Groups[1]

Analúcia D. Schliemann, ORCID 0000-0001-6365-1466
Department of Education, Tufts University, Medford, MA, USA[2]
aschliem@tufts.edu, Phone 1-978-852-6067.

Mirjana Hotomski, ORCID 0000-0002-3968-8565
Department of Education, Tufts University, Medford, MA, USA
mirjana.hotomski@tufts.edu, Phone 617-645-8047.

Montserrat Teixidor-i-Bigas, ORCID 0000-0003-3747-3330
Department of Education, Tufts University, Medford, MA, USA
montserrat.teixidoribigas@tufts.edu, Phone 1-781-979-9161

David W. Carraher, ORCID 0000-0003-2356-9438
TERC, Cambridge, MA, USA[3]
david_carraher@mac.com, Phone 1-978-852-6067

Correspondence:
Analúcia D. Schliemann, Professor Emerita
Tufts University, Department of Education, Paige Hall, Medford, MA 02155
aschliem@tufts.edu, Phone 1-978-8526067

---

[1] This study was developed as part of the National Science Foundation grant # 0962863. Opinions, conclusions, and recommendations are those of the authors and do not necessarily reflect the Foundation's views.

[2] Professor Emerita, Retired.

[3] Senior Scientist, Retired.

Teacher Development and Mathematics Performance Across Ethnic Groups

Abstract

We describe the implementation and assess the impact of a mathematics teacher development program that uses variables, functions, and multiple representations to integrate topics in the mathematics curriculum. Examples of lessons by teachers and their survey answers suggest improvements in teaching and learning in the target multi-ethnic district. Student performance on state-mandated assessments of mathematics for grades 3 to 5, before and after teachers' first year in the program, show that the gains in proficiency for the target district were more than three times as great as those for comparison districts and for the state as a whole. This ratio of gains held for the African-American students. Gains for Hispanic/Latino students were more than four times as those for comparison districts and more than three times the gains for state districts. Gains for African-American and Hispanic/Latino students in the target district were larger than those for White students, with a slight decrease in achievement differences between African-American and White students and a relatively large decrease in the difference between Hispanic/Latinos and White students. In the comparison and state districts gains for White students were larger than gains for the other two groups. The mean value of the assessment's five levels also showed significant gains in the target district, where African-Americans improved as much as White students and Hispanic/Latino students showed even larger gains. We discuss program features and changes in teaching that may have contributed to better learning among target district students across the three ethnic groups investigated.

Introduction

This study examines the potential contribution of a teacher development program to the teaching and learning of mathematics in a multi-ethnic district. The program uses functions and their multiple representations (e.g., natural language, line segments, function tables, Cartesian graphs, and algebraic notation) in ways designed to integrate topics in the mathematics curriculum, from arithmetic to algebra, thus aiming to promote mathematical reasoning and a deep understanding of mathematics through analyses and discussions about relations involving numbers and quantities.

The view that reasoning about functions has a major role to play in the teaching of algebra in secondary mathematics education has often been embraced by mathematics educators (e.g. Harel and Dubinsky, 1992, Oehrtman, Carlson, & Thompson, 2008, Schwartz & Yerushalmy, 1992, and Seldon & Seldon, 1992). Moreover, mathematics educators participating in the U.S. Department of Education's Algebra Initiative Colloquium (LaCampagne, Blair, & Kaput, 1995), have argued that, instead of appearing in isolated courses in middle or high school, algebra should pervade the K-12 curriculum to lend coherence, depth, and power to school mathematics from the first years of elementary school, thus preparing students for later high school algebra courses (e.g., Kaput, 1995, 1998; Schoenfeld, 1995). A number of classroom studies of algebra in the early grades have adopted an approach to algebra in which functions are given a major role. Results have shown that, starting from Kindergarten, very young students can learn to successfully reason and represent variables and function relations (e.g., Blanton, Brizuela, et al., 2015, Blanton, Stephens, et al., 2015, Carraher & Schliemann, 2018). By grades 4 and 5 (see Carraher, Schliemann, and Schwartz, 2008, for details of the study), students who had participated, from grade 3, in weekly early algebra lessons within a particular functions approach, represented verbal problems as graphs and as equations and used the Cartesian space as well as the syntactic rules of algebra to solve equations. They did so as they were practicing number facts, reasoning about arithmetic operations properties, and developing a broader view of arithmetic, of relations among sets of quantities and numbers, and of the geometrical representation of arithmetic operations. Later data from a sample of students who had participated in that intervention study suggest that classroom activities on algebra and functions contributed to their better performance in mathematics in grades 7 and 8 (Schliemann, Carraher, & Brizuela, 2012).

The particular elementary and middle school activities regarding algebra and functions carried out by researchers have yet to be brought into the regular practice of school teachers, even though the National Council of Teachers of Mathematics (NCTM, 2000) and the Common Core State Standards Initiative (CCSSI, 2010) call for inclusion of a K-12 algebra strand and, in the case of NCTM, for a focus on variables and functions from the early school years (see discussion by Carraher & Schliemann, 2019).

In view of NCTM (2000) recommendations and of results of classroom research, it would appear to make sense to prepare teachers, coaches, and special educators to work, throughout the school years, with concepts and representations related to algebra, variables, and functions. This would require providing teachers with an integrated vision of the mathematics content they teach and of how they teach, through the lens of functions. Such a substantial change requires long-term teacher development programs capable of integrating mathematics content and teaching responsive to students' reasoning and initial representations. In addition, researchers should evaluate how this preparation may (or may not) lead to advances in teaching and student learning.

The present study evaluates the contribution of a teacher development program aimed at preparing teachers to use algebra, variables, functions, and their multiple representations as integral components in their coverage of standard topics in grades 3 to 8.

An external evaluation of lessons taught by teachers in the first program's cohort, using the Reformed Teaching Observation Protocol (RTOP, AZ State

University, Sawada et al., 2002), showed that, from the start to the end of the program, students in the nine participating districts, all in Massachusetts, more often engaged in discussions, put forth their own ideas, considered each other's ideas, and showed interest, motivation, and on-task behavior. At the end of the courses, the mean RTOP score had increased from 43.1 to 49.1, with an effect size of .30 ($p<.05$). One year after the first cohort of teachers had completed the program, the mean RTOP scores ratings further increased to 57.5, with substantial effect size of 0.69. When the program began (Spring of 2011), the percentage of students at the Proficient and Advanced levels at the then mandated state assessment (Massachusetts Comprehensive Assessment System - MCAS) was nearly identical to that of similar comparison districts (47.9% vs. 48.0%). Three years later, the target districts had surpassed comparison districts by 4.5% and had narrowed the gap with regard to the state from 7.1% to 4.5%. Details of these analyses are found in Teixidor, Schliemann, & Carraher (2013), Schliemann, Carraher, & Teixidor (2016), and https://sites.tufts.edu/poincare/research-and-impact/).

In the present study we focus on data from a fourth cohort of teachers in a single Massachusetts target district. We illustrate changes in teaching and learning through examples of lessons taught at the start and end of the program, as well as examples of teachers' answers to anonymous surveys. We further evaluate the contribution of the program to student learning focusing on changes among African-American, Hispanic-Latino, and White students' performance. We compare changes in the target district to changes in ten similar districts and in the whole state.

The Program's Foundations and Research Goals

The program embraced the idea that elementary and middle school students can benefit from learning to express relations between variables before they are formally introduced to functions or encounter functions in equations, where variables are "unknowns" constrained to one or two values, namely, solutions. Functions are arguably already part of the school curriculum before they are named as such or rigorously defined. What we are proposing is that functions should have a different sort of presence and role than they currently do. For example, “four plus”, and “two times” (commonly represented as $x + 4$ and $2x$) can be regarded as functions (Gowers, Barrow-Green & Leader, 2010, p. 10). Within a broader framework, in which functions are assignments from a certain domain (not necessarily the real line) to a given codomain, addition is a function with pairs of numbers as inputs and a single number as output.

In focusing on functions as a core concept across the K-12 curriculum, we are not proposing to expand the curriculum or to ask students to memorize or be familiar with the definition of functions. Instead, we propose that teachers be prepared to use variables and functions to organize and integrate different topics within and across the school years, highlighting the fact that these topics, traditionally taught in isolation, are in fact interrelated. In our view, this integration holds the promise of helping teachers and their students understand in greater depth the meaning and use of multiple curriculum concepts by:

- Allowing students to work on open-ended questions that elicit generalizations about relations involving sets of numbers or measures and variables, instead of focusing on computations on specific values.
- Engaging students in producing their own representations for problem situations and subsequent discussion and introduction of new representations.
- Using multiple representations, including drawings, verbal description, number lines, data tables, function graphs, and algebraic notation for relationships between sets of numbers or quantities.

- Integrating mathematics topics across the school years, from arithmetic operations to fractions, ratios, proportions, slope, linear equations, quadratic equations, and change and invariance, through their common connections in the language of functions.
- Helping students interpret equations as a comparison of two functions.

The program was developed by a team of mathematicians, physicists, and mathematics education researchers seeking to create an integrated approach between the mathematical knowledge in the school curriculum and a teaching approach that focus on students' reasoning. Online course materials included extensive notes on mathematical content, problems to discuss and solve, interactive software demonstrations, and classroom teaching examples for discussion and adaptation. Online and face-to-face activities aimed at mentoring teachers through discussion of mathematical content and their classroom implementation activities. Details on the program can be found at (https://sites.tufts.edu/poincare/).

Features of the program matched what Hiebert and Grouws (2007) identified as two crucial features that would facilitate student conceptual understanding and perhaps even skill fluency. One factor is explicit attention to connections among ideas, facts, and procedures; the other is student engagement in struggle with important mathematics topics that are within their reach but not yet fully developed. Program courses strongly focus on teachers' mathematical knowledge and engage participating school districts in supporting program activities, two factors that, as Hill, Blazer, and Lynch (2015) found, explained a moderate share of the variation in mathematics-specific teaching dimensions in fourth- and fifth-grade classrooms from the four school districts they studied. The program also aims at intellectual engagement, interactions between students and teachers, accuracy with regard to content, meaning-focused instruction (as opposed to rote learning), close attention to students' perspectives, focus on questioning and discussion, and engagement of students in classroom participation. These are qualities of effective teachers found in Kane and Staiger's (2012) extensive study of nearly 3000 volunteer teachers and students.

From 2010 to 2016 the three courses in the program were offered to four consecutive cohorts, each of approximately 60 teachers. We report here on the fourth cohort of teachers, a group that included 60% of the teachers, mathematics coaches, interventionists, and special education teachers in grades 3 through 10, from a single multi-ethnic and low-income target district. After describing the program, we present (a) an exploratory qualitative analysis of examples of teachers' classroom lessons and corresponding student participation, at the start and at the end of the program, (b) examples of participating teachers' answers to anonymous surveys taken at the end of courses, and (c) a quantitative analysis of changes in state mandated assessment results of African-American, Hispanic/Latino, and White students in grades 5 though 8, from the start to the end of the first year the teachers took two of the program courses; we compare changes in the target district to changes in districts with similar populations and similar initial assessments results and in the whole state. In light of the findings, we will discuss program features and improvement in teaching that may have contributed to better learning among students in the target district.

Research questions and related studies

We consider the following questions:

(a) Would teachers in the program change the way they plan and implement their classroom lessons while incorporating a functions approach to their teaching?

(b) Would the functions approach to mathematics promoted by the program lead to changes in classroom discussions and in teachers' facilitation of and attention to student reasoning?

(c) Would students of teachers in the program, even in the earlier grades, successfully reason about, interpret, and represent variables and function relations?

(d) Does the particular program's approach to mathematics help students learn the mathematics addressed in the school curriculum and evaluated by mandated standard assessments, which are not related to variables and functions?

(e) Does the program's approach benefit students from different ethnic backgrounds?

The exploratory analysis of examples of teachers' lessons and teachers' answers to anonymous surveys may serve to identify some of the changes in teachers' practice and in students' participation in classroom activities that included variables, functions, and their representations. This should provide tentative answers to questions a, b, and c and will help in identifying factors that may have contributed to student performance on state-mandated standard assessments. The quantitative analysis of student assessment performance addresses questions d and e and constitute the main findings of our study.

First, let us broadly consider theoretical foundations and research studies relevant to our approach and to each of our research questions.

On question (a) Would teachers in the program change the way they plan and implement their classroom lessons while incorporating a functions approach to their teaching?

Chazan (1999, 2000) reports that making functions central to an algebra course helped his students solve problems before learning standard methods to solve equations.

Courses on the concept of function specifically tailored for teachers have been developed and evaluated in terms of teachers' learning. For example, Steele and Hillen (2012) and Steele, Hillen and Smith (2013) developed a course whose goal was to teach the definition of function to a mixed group of teachers. They found that, by the end of the course, teachers better understood the concept of function, could distinguish between examples and non-examples of functions, and favored activities that would encourage the use of the notion of function in the classroom. However, these courses aimed at teaching directly about functions per se, rather than preparing teachers to work on a broad range of mathematics topics, from arithmetic to algebra, through the lens of functions, as was the case in our program.

To answer question *a*, we describe examples of how teachers incorporate variables, functions, and multiple representations into the teaching of topics in the elementary and middle school curriculum, a distinctive characteristic of our program. Teachers' answers to surveys on the courses they took further clarify the program's contribution to their teaching practice.

On question (b): Would the functions approach promoted by the program lead to changes in classroom discussions and in teachers' facilitation of and attention to student reasoning?

Piaget's and Vygotsky's ideas are the basis of socio-constructivist views of education and of many programs intended at improving the teaching and learning of mathematics. Their work also provides a general framework to our program.

Piaget's theory and empirical analyses stress the relevance of children's own construction processes towards understanding of logic, physics, and mathematics concepts. Piaget (1995) also highlighted that social interactions based on the transmission of ready-made representations or procedures only provide the individual with superficial notions; cooperation, on the other hand, described as a system of interpersonal actions, allows for the development of a coherent system of knowledge, as participants use similar systems of representations and meanings and carry out similar operations. Giving emphasis to cultural contributions to cognitive development and learning, Vygotsky (1978) proposed that representational tools mediate thought processes and transform cognition, channeling and structuring thinking in new ways. The tools and representational systems the individual has access to and uses play a role in how mathematical thinking is structured, allowing for new different aspects of mathematical relationships to come to the forefront.

Reasoning about physical quantities to represent the mathematical relationships among them is an important source of mathematical learning. While students may not be capable of approaching mathematical ideas presented in an abstract way, they are able to explore and discuss numerical or quantitative relations presented in a contextualized framework, using their own resources and intuitive representations. In these settings, new forms of representation can be introduced, further deepening mathematical understanding. These aspects match Piaget and Vygotsky's constructs and have been supported by a wide range of research, including Piaget and Inhelder's (1974) investigations of children's reasoning about quantities and by studies of the role of contexts and everyday situations in the development of mathematical knowledge and reasoning (Carraher, Carraher, & Schliemann, 1985; Nunes, Schliemann, & Carraher, 1993; Reed & Lave, 1979; Saxe, 1991), quantitative reasoning (Smith & Thompson 2008; Thompson 2011), and the realistic mathematics education approach (Freudenthal, 1973, 1991; Gravemeijer, 1999). Specifically, research has shown that reasoning about relationships among physical quantities leads to conceptual understanding of powerful forms of representation and understanding of mathematical procedures (e.g.

Freudenthal, 1973, 1991; Gravemeijer, 1999; Kaput, 1995; Kirshner, 2001; Lobato & Ellis, 2002; Olive & Caglayan, 2008; Smith & Thompson, 2008; Thompson, 2015).

Because functions can be represented through non-algebraic forms (for example, natural language or drawings), students can work with relations corresponding to equations before they have mastered algebraic notation (see, for example, Brizuela, Blanton et al., 2015, and Carraher & Schliemann, 2018 review of studies of early algebra). The process of moving from the specific to the general by incorporating variables, from arithmetic to algebra, lies at the core of our functions approach to teaching mathematics. The teacher can then help students compare and discuss the virtues and drawbacks of various representations they may produce and introduce new mathematical representations and new tools along the way.

In addressing question *b*, we describe examples of how participants in the teacher development program incorporated activities and representations of functions in their teaching, as they implemented activities where students considered relationships among physical quantities and expressed their views and representations, further expanding these through classroom discussion and access to new concepts and representations.

On question (c): Would students of teachers in the program, even in the earlier grades, give evidence of reasoning about, interpreting, and representing variables and function relations?

It would seem that questions on whether *students* actually benefit from their teachers' participation in professional development programs should be a major focus of research on teacher development. Nevertheless, as pointed out by Gersten et al. (2014), Hill (2007), Bautista (2015), Bautista and Ortega-Ruiz (2015), Sztajn et al. (2017), and others, for many decades remarkably little research has tried to evaluate and uncover evidence of an impact of professional development on student achievement. Recently, however, as Darling-Hammond, Hyler, and Gardner (2017) conclude from their extensive analysis of teacher development programs in various subject areas, studies by Akiba & Liang (2016), Desimone et al. (2013), Franke et al. (2001), McMeeking et al. (2012), Santagata et al. (2011), and Saxe et al. (2001) constitute evidence that changes in teacher's knowledge and practices through professional development are associated with improvement in teaching and in student learning of the mathematics in the school curriculum. From these programs' description, they further claim that successful teacher development programs include content focus, active learning, collaboration in the context of teaching, use of examples of effective practice, coaching and expert support, opportunities for feedback and reflection, and sustained duration.

Others have considered similar aspects as critical for promoting learning for all students, including those underperforming on standard assessments, such as conceptual learning, in-class time spent on teaching, teachers' analyses of students' difficulties and strategies, attention to students' ideas, classroom discussions, students' interactions while answering open ended questions, activities on measurement and geometry, and teacher collaboration (e.g., Boaler & Staples, 2008; Brown, 2012; Desimone & Long, 2010; Fernandes, Crespo, & Civil, 2017). Singham (2003) emphasized that teacher development programs that would make a difference in teaching and learning are not "scattershot, single-session, workshop-style programs that pass for professional development in so many school districts" (p. 590), but rather constitute sustained programs combining both mathematical content and specific issues of classroom practice and interactions.

Given research findings regarding algebra and functions in the early grades (Blanton, Stephens, et al., 2015; Carraher & Schliemann, 2018), we believe that most of the above aspects contributing to student learning can be facilitated and expanded by our focus on functions and their representations. But would the elementary and middle school teachers help their students understand and use, implicitly or explicitly, variable, functions, and their multiple representations (question *c*)? To address this question we examined student participation in examples of lessons implemented by teachers in the program, focusing on students' use and representation of variables and functions.

On question (d): Does the program's approach to mathematics help students learn the mathematics addressed in the school curriculum and evaluated by mandated standard assessments, which are not related to variables and functions?

Traditionally, student achievement throughout the US is evaluated by standardized state-mandated assessments, widely viewed as limited in failing to detect conceptual reasoning, being unfair to students from different ethnic groups and language background, and/or leading teachers to teaching to the test (see ProCon.org, n.d., for a list of pros and cons of standardized tests and supporting references). Nevertheless, assessment scores constitute one of the few available measurement tools to detect differences in student achievement regarding curricula adopted by schools across different ethnic populations.

Using standard assessment scores, Riordan and Noyce (2001), found that grade 4 students in 88 Massachusetts schools implementing mathematics reform curricula aligned with the NCTM (1989) outperformed their peers in schools matched according to previous assessment scores and with similar percentage of students with free or reduced-price lunch, but teaching according to traditional curricula. This was true for White, non-White, and low-income students. However, as the authors point out, students in their study were predominantly White, performed above the state average on standard assessments, and only a small percentage of them was eligible for free and subsidized lunches.

To answer question *d*, we examined grades 3-8 student performance on state mandated standardized assessments in a low-income district, with large percentages of African-American and Hispanic/Latino students, and relatively low scores on standard assessments. Despite their limitations, the fact that the content of the assessments was not directly related to the development program offered to the teachers, their results allow for examining the program's effectiveness in improving mathematical knowledge across the state curriculum topics. In this study, standard assessment scores constituted an independent measure for comparing different groups and districts' results and for evaluating progress over the year, for a relatively large number of students.

On question (e): Does the particular program's approach benefit students from different ethnic backgrounds?

One would hope that efforts to improve teaching should benefit students, regardless of their ethnic background. This, however, does not seem to be the case. For example, Desmond-Helmann (2016), reporting on the Gates Foundation's efforts to improve teaching and learning in Kentucky, the first state adopting the Common Core standards, notes that, during the period of the Foundation support (2010-2016), students meeting three out of four ACT College Readiness Benchmarks increased from 27 percent to 33 percent. Across the country the corresponding percentages did not change. Nevertheless, as reported by the Lexington Herald Leader (https://www.kentucky.com/news/local/education/article97527057.html), although 2016 scores for Kentucky's public school graduates improved from where they were years earlier, racial achievement differences persisted, with composite scores (the average of English, mathematics, reading, and science scores, ranging from 1 to 36) of 20.3 for

Whites and 16.8 for African-American students. In addition, from 2012 to 2015, even though the percentage of elementary school students proficient in mathematics in Kentucky had increased from 40 to 49 %, only 31% of African-American students were deemed “proficient” compared to 52% of White students. For a fixed group of students, this difference increased as they progressed from one grade to the next (see Ostashevsky, 2019).

In the past, ethnic group differences in academic achievement have often been attributed to cultural deprivation and disadvantages of underperforming groups (Bloom, Davis, & Hess, 1965; Brooks-Gunn & Duncan, 1997; Coleman, 1966; Duncan et al., 1998). Many, including teachers (Bol & Berry, 2005), have attributed poor performance to shortcomings in students’ motivation, work ethic, or lack of support from family members. However, recent research has shown that teachers of students of low socio-economic status tend to be weaker in mathematics and pedagogical knowledge (Bol & Berry, 2005; Hill & Lubienski, 2007). Hill and Lubienski (2007) observed that “the percentage of free-lunch eligible students in a school is significantly related to teachers’ […] scores” on a test of mathematical knowledge and to their qualifications (p. 761). This shift of attention to quality of teaching led to recommendations for improving learning among underperforming student groups through culturally responsive teaching, outreach to families, more funding, longer school days, and, most important, investing in teacher quality and teacher development (http://www.nea.org/home/13550.htm, Akiba et al., 2007, Hirsh, 2005; Ladson-Billings, 2000; Schoenfeld, 2002). As Duncan and Murnane (2014a) summarizes:

> Discussions of school reforms often center on simplistic silver bullets: more money, more accountability, more choice, new organizational structures. None of these reforms has turned the tide because none focuses directly on improving what matters most in education: the quality and consistency of the instruction and experiences offered to students (p. 14).

Gutierrez (2008), in a critique of the excessive focus on reporting the so-called achievement gap between different ethnic groups, calls for less research on documenting achievement differences, causes of the differences, or single variables that predict success, proposing, instead, the development of more research on effective teaching and learning for underperforming groups, along with rich descriptions of these environments, as well as intervention studies, including professional development. A few intervention studies along these lines were successful in contributing to better learning among minority students and in identifying factors that may have contributed to their results.

Duncan and Murnane (2014b, c) report on educational interventions in schools in three urban areas (Boston, Chicago, and New York) were associated with better performance for students from low-income minorities. They summarize their findings as follows:

> All of them take advantage of *advances in research knowledge* about the active ingredients of good pre-K, elementary, or high school education. All provide important *school supports* for teachers and school leaders. All incorporate *sensible systems of accountability*. And finally, all incorporate *high academic standards. (Duncan & Murnane, 2014b, p. 51*).

Boaler and Staples (2008) found that, in comparison to schools using traditional methods, students in a target school showed significantly higher levels of achievement, with reduction of differences in attainment between different ethnic groups. They attribute their positive results to multiple factors:

> The discussions at Railside were often abstract mathematical discussions and the students *did not learn mathematics through special materials that were sensitive to issues of gender, culture, or class*. But through their mathematical work, the Railside students learned to appreciate the different ways that students saw mathematics problems and learned to value the contribution of different methods,

> perspectives, representations, partial ideas and even incorrect ideas as they worked to solve problems. As the classrooms became more multidimensional, students learned to appreciate and value the insights of a wider group of students from different cultures and circumstances (p. 639-640).

Briars and Resnick (2000, see also Schoenfeld, 2002) describe a three-year sustained implementation of a reform curriculum and teacher development in Pittsburg, PA schools. The teacher development program, built around the NCTM (2000) standards and their implementation, aimed at coherent and connected mathematical content and high expectations for all students in reasoning, representation, communication, problem solving, and making connections. They found that the implementation led to improved assessment results for African-American fourth graders, closer to those of White students in the district.

To answer question *e*, we evaluate the potential impact of our program on learning by different ethnic groups by examining changes in student performance on standard assessments. In doing so, we used the same yardstick that has been traditionally employed to determine differences in performance across ethnic groups.

In our discussion section, we attempt to identify the characteristics of our functions approach that may have contributed to improved learning for the target district students across the different groups.

Program Activities

The program was designed to (a) offer teachers access to a new view of key topics in the curriculum; (b) prepare them to elicit, discuss, and raise new questions about students' initial ways of reasoning and representing problem situations; (c) promote discussions among students; and (d) introduce new tools for representing mathematical ideas and for solving problems.

Course structure

Functions relations, variables, and their representations are to serve as a unifying thread across topics in the elementary, middle, and high school curricula. Teachers explore arithmetic operations, fractions, ratios, proportions, algebra, and geometry from the perspective of functions and jointly discuss and solve open-ended problems. Course activities aim at generalizations and multiple representations for relationships among sets of numbers or physical quantities through language, number-lines, graphs, and mathematical notation. These characteristics of the courses were intended to integrate mathematical content and pedagogical knowledge in the classroom.

The semester-long courses were hosted online and complemented by routine face-to-face meetings among teachers at school sites. Each course consisted of four units and a final project. Two sets of resources of different levels of complexity were available, one for teachers of grades 3 to 6, the other for those in grades 5 to 10. Early middle school teachers could select the set that would better fit their mathematical background. The teachers invested, on average, 10 hours of work per week on program activities.

The first two weeks of each unit focused on texts on mathematics content or on findings from education research, video-lectures, software demonstrations, open-ended problem situations, and videotaped classroom lessons. The software demonstrations, specifically designed by members of the program's team, allowed teachers to explore selected mathematical concepts. The videotaped classroom lessons were analyzed by groups of teachers in terms of teaching and of students' reasoning. Online groups of 8 to 10 teachers discussed the texts and video materials and solved the problems. In the third week of each unit, teachers worked in small groups of 2 to 4 members from the same school, but generally not from the same online group, interviewing students about particular topics, and planning, implementing, and evaluating classroom activities. At the end of each course, teachers in these small school groups worked together in designing a lesson on a topic of their choice. The lesson was implemented, with adaptations, by each teacher in the group. The teachers then jointly analyzed students' responses and participation in the lessons. Following this, each teacher produced an individual report on the implementation and on ways to improve the lesson so as to further students' learning.

Each online and small school group had a mentor (a faculty member, researcher, postdoc, graduate student in mathematics education, or a selected teacher from a previous cohort), who read daily posts by teachers and responded to teachers' answers, raised questions to trigger discussions and further reasoning, and provided suggestions and feedback on work underway. The discussions and feedback were meant to lead to improved problem solving work and better lesson plans. Mentors also held monthly face-to-face meetings with their groups.

The content of the courses

The course content was generally aligned with the topics discussed in elementary, middle, and beginning of high school mathematics courses, but did not make reference to the particular textbooks used by the schools or to the state assessments. The mathematics focus fell mainly on foundational and structural mathematical concepts, rather than on procedural steps.

The first course introduced functions and their expression through verbal, tabular, graphical, and algebraic representations. In effect, functions were offered as a lens for investigating numbers (including fractions) and arithmetic operations. Variables were used to highlight relations among quantities and to support the formulation of mathematical generalizations.

The second course focused on the idea that equations and inequalities could be construed as entailing comparisons of two functions. Equations were represented algebraically and through the graphs of the associated functions (each corresponding to the terms on one side of the equation). Teachers reviewed how solutions to equations correspond to the x-coordinates of points of intersection of the graphs of two functions. Transformations were introduced as mappings of a set onto itself and expressed either as a line in the plane or a plane in 3-space. Such transformations were expressly linked to the operations of addition and multiplication. Steps in equation-solving were treated as (a) the application of the same transformation on two functions or (b) as the application of the same transformation to each graph in the plane. Divisibility of integers was associated with divisibility for polynomials and the solution of polynomial equations.

The third course dealt with systems of linear equations and with the concepts of change and invariance in the case of linear and non-linear functions.

The classroom videos included in course materials and discussed by participating teachers were selected from a collection that had been analyzed by early algebra researchers (see Carraher & Schliemann, 2016, 2018). The videos feature students, from grades 3 to 7, discussing and making generalizations, using variable notation and graphs of linear functions. Teachers used the videos as points of departure for planning their own classroom activities. The research videos presented a variety of situations in which students were encouraged to employ new representations to give expression to their own observations. The students also used their own drawings, number line diagrams, data tables, graphs, algebraic expressions, and equations as models of relations among physical quantities. In this process, the individual students' contributions were considered and discussed by their peers and by the teacher.

Method

In the target district, situated in the Greater Boston Area, 18% of the students were African-American and 44% were of Hispanic/Latino origin. They were taught mathematics by 83 teachers and received direct or indirect input from other 21 professionals assigned to their classrooms. Sixty three of the educators (61%) took the first two courses in the program from August 28, 2015 to December 11, 2016. The group was comprised of 53 elementary, middle, and high school teachers of mathematics and ten special education teachers, coaches, or interventionists.

Teachers volunteered to enroll in the courses after an hour-long information session and discussion headed by program leaders and encouragement by the district's curriculum coordinators. Many demonstrated enthusiasm about participating, even after they were made aware of its high demands. Over the three semesters, a grant provided the teachers with stipends, computers, and tuition.

Teachers and administrators in the target district were concerned about addressing the recently implemented Common Core Standards, adopting new textbooks, and looking for teacher development options for all subject areas. Teachers also participated in short bi-weekly or monthly teacher development meetings on general educational and administrative issues, as commonly found across school districts.

Because participants were volunteers, they were not a random sample. This fact might have somehow benefitted the achievement of their own students. However, the student data included the assessment results of all students in the district, not simply the students of the participating teachers. This would appear to control for teacher selection bias issues that might have otherwise arisen.

The performance of target district students in state mandated assessments was compared to that of students from ten similar school districts. The comparison districts were chosen for having (a) similar percentages of students at the proficient levels of performance in the state mandated assessment at the end of school year 2014-2015 and (b) similar characteristics of teachers and student population. We also compared the target district's assessment results to those of the whole state.

On changes in teaching and student classroom participation

To illustrate changes in teachers' ideas and teaching (research questions *a* and *b*), we examined the work of two teachers, one at the elementary school, the other at the middle school level, at the start of course 1 (unit 2), while they were implementing an activity proposed by the program, and at the end of course 3 (final project), as they were implementing lessons designed in collaboration with their peers. These implementations were also analyzed in terms of students' ideas, representations, and contributions to classroom discussions during activities implemented by the two teachers, at the start and at the end of the program (research question c). The two teachers were among those judged as presenting weak classroom implementations at the start of the program. This analysis is complemented by an account of teachers answers to written surveys.

On changes in student assessment performance

To answer research questions *d* and *e,* we examined changes in student results in grades 3-8 on the state mandated assessment, from before teachers entered the program to a year later, when teachers were finishing the second course. At the time, the state of Massachusetts had adopted the assessment developed by the Partnership for Assessment of Readiness for College and Careers scores (PARCC, see http://profiles.doe.mass.edu/state_report/parcc.aspx). Data for grades 9-10 are not included in the present study because students were not tested in grade 9 and different tests were used across schools in grade 10.

Issues of confidentiality prevented us from obtaining individual student reports and to analyze how their results related to their teachers' performance in course activities. Instead we compared the changes in performance of students in the target districts to those of students from non-participating, similar districts—similar in terms of demographics and past performance on state assessments. Any significant difference between gains in the targeted district versus gains in the similar comparison districts could arguably be taken as evidence of the impact of the program on students' mathematical achievement[4].

Timeline and districts data

The timeline of courses and assessments was as follows:

- PARCC assessment before courses: from May 4 to May 29, 2015.
- Course 1: from August 28 to December 14, 2015
- Course 2: from January 4 to May 6, 2016
- PARCC assessment after courses 1 and 2: from April 25 to June 6, 2016.

It was not possible to perfectly match target and comparison districts on each measure. Nonetheless, the intervention and comparison districts were generally similar in terms of several characteristics of students and teachers (see Table 1).

Table 1: Student and teacher characteristics at the outset of the study

Students' Characteristics	Target	Comparison	State*
Initial Percentage Proficient (2014-2015)	37.0	35.6	49.1

[4] In the discussion section we weight whether alternative interpretations might be given for the significant results obtained.

Initial Average math performance (2014-2015)	3.009	2.973	3.285
Percentage Minority Students	69.2	47.4	37.3
Percentage First Language not English	58.6	28.7	19.0
Percentage English Language Learner	16.0	13.1	9.0
Percentage Economically Disadvantaged	42.1	39.7	27.4
Percentage with Disabilities	15.1	17.9	17.2
Percentage High Needs	61.9	54.5	43.5
Number of students in 3-8 grade taking PARCC	2,769	34,557	225,579
Students' family median income	$50,762	$54,795	$68,563
Percentage of African-American students	18%	14%	9%
Percentage of Hispanic/Latino students	44%	25%	19%
Percentage of White students,	31%	52%	63%

Teachers' Characteristics	Target	Comparison	State
Licensed in Teaching Assignment	100	98.6	97.4
Classes with Highly Qualified	99	97.4	96.3
Student Teacher Ratio	13.6 to 1	13.9 to 1	13.2 to 1
Average Salary	$76,262	$72,899	$74,782
% Retained	82.1 %	85.0 %	85.5 %
Exemplary	3.3 %	10.3 %	10.7 %
Proficient	91.1 %	82.5 %	84.7 %

*State data refer to districts taking PARCC Assessments in both 2014-2015 and 2015-2016 academic years.

Results

Changes in teaching and in student classroom participation from the start to the end of the program (questions a, b, and c)

The following qualitative description of the classroom activities by two teachers, at the start and at the end of the program, aims at tentatively answer research questions a, b, and c. The analysis was first carried out by one of the authors of this paper and then checked and, if necessary, changed and improved, by two of the other authors.

An activity at the start of the program

In the fourth week of the first course, the teachers observed and analyzed a video from classroom research where the instructor started a lesson (the Candy Boxes lesson) by holding two opaque boxes and informed the students that (a) each box has the same number of candies, (b) one box is John's, and it contains all of his candies, and (c) the other box, along with three additional candies, belongs to Mary. In the research video example, when the instructor asked the students to represent the problem on paper, many students assigned specific values to each amount. After the instructor drew a data table on the board with columns for number of candies in a box (the number John could have) and number of candies Mary's would then have. Each student suggested a value for the number of candies in the box and computed the number of Mary's candies. Inconsistent values were eliminated because, as one of the students expressed it, "Mary has to have 3 more than John." The focus of the lesson then shifted from computations on specific pairs of numbers to determining a general rule (n, n+3) that applies to all ordered pairs in the problem. This allowed for discussing variables as placeholders for any possible number of candies in each box and the introduction of a functional relation, such as $n \rightarrow n+3$, that is, the application of the function, $f(n) = n+3$.

After watching and discussing the video, the teachers were to plan and implement an activity (interview with a few students or a full classroom lesson), adapted from what they had seen in the video. The instructions stated that:

> "… even advanced students tend to interpret letters as just a fixed unknown number to be found. With tasks similar to the Candy Boxes, the goal is to help

students consider variables as representing any possible value for the amount they stand for."

The teachers were asked to present and discuss the problem situation, making sure students understood it. They were to have the student (a) represent on paper John's and Mary's amounts, (b) discuss their representations, (c) fill in a table listing John's and Mary's corresponding possible amounts, and (d) attempt to make generalizations regarding the relationship between the two sets of amounts in the table. If no student proposed to employ a letter to stand for the variable, the teacher was to suggest doing so by inserting a letter in the last row of the column for John's number of candies (the number of candies in the box). She then asked students to propose and discuss how to show that Mary has 3 more candies than John, regardless of the number of candies in the box.

An activity at the end of the program

For the Final Project in Course 3 (last course in the program) the teachers in each school group jointly planned a lesson on a topic of their choice, individually implemented the lesson in their classrooms, discussed the lesson implementation with teachers in the group, and individually reported on the implementation and on possible changes for improving the lesson.

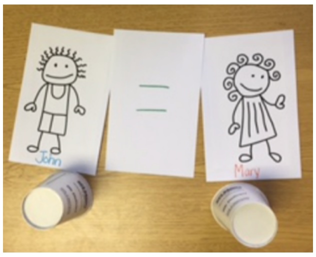
In the instructions for the Final Project lesson, teachers were asked to design an activity that (a) would engage students in exploring new concepts and new mathematical ideas and representations, (b) avoid to focus on algorithms, rules to be memorized, or starting with a request for a definition of a term or a straightforward computation, (c) focus on an open-ended problem for which students are likely to come up with different approaches and put forth various, possibly conflicting, ideas and representations that would lead to discussions and, hopefully, to new views and representations.

Teachers video-recorded the lessons and transcribed what they deemed the most interesting moments in their reports.

The elementary classroom activity at the start of the program

One of the two teachers in our analysis worked with two first graders. She worked with blocks, instead of candies, and started by making sure the students understood that Mary’s and John’s containers would have equal amounts of blocks. She showed cards with the symbols +, -, and =, and asked the students to use them to represent that the two boxes had the same amount. After a student suggested to use the equals sign, the teacher illustrated the equality using pre-prepared cards (see Figure 1). She did not encourage the students to come up with their own representation for the problem and, contrary to instructions, introduced her own drawing for the equality.

Figure 1: The representation that John’s amount was equal to Mary’s amount, prepared by the teacher



After that the teacher placed three cubes on top of Mary’s box and asked the students: “*What would happen if I placed 3 cubes onto Mary’s box? How many cubes are there now?* A student answered that “*If there are three over here in John’s and three over here in Mary’s then three plus three is six.”* The other student said that *“Mary has 6 cubes and John will have 3.”*

The teacher then focused on other possible amounts in the box by asking “*what if there wasn’t three cubes in the boxes?”* A student answered that there could be 5 blocks in the box and another says that if there were 5 blocks in the box Mary would have 7 (not 8) blocks.

Next, instead of registering values in a table and asking students to formulate generalizations, the teacher herself stated that Mary would always have three more and the students agreed with her. Only then did she introduced a data table and proceeded by asking students to compute Mary's total amounts for a few possible amounts of blocks in the box.

The analysis of her implementation revealed that, six weeks into the program, the teacher did not build upon the students' own representations of the situation nor did she employ a table relating John's and Mary's number of blocks. Instead, she announced the generalization herself and used the table to ask students to perform computations. She elicited answers by individual students, but did not encourage discussions between them.

The elementary school activity at the end of the program

In the following illustrative example, the same elementary school teacher who taught the Candy Boxes lesson in the example above, implemented her final project lesson with eight third grade students. The teachers in her elementary school group had been working with units of measurement and conversions from inches to feet. This group's initial plan aimed at helping "*students to gain a better understanding of unit conversions involving measurement.*" After discussions and feedback from the group's instructor, the group decided to introduce an unconventional measuring unit, the length of a domino piece and start the lesson by eliciting students' ideas about measurement. They proposed that one group of students would measure the length of various paper strips with dominos while the other would measure the strips with rulers, recording the results in a table.

The group of students the teacher in our example worked with included both special education students and one English Language Learner. The teacher began eliciting students' ideas about what could be used for measuring length and how to measure with different tools. The students mentioned measuring tapes, rulers, yardsticks, and meter sticks. A student demonstrated how to use a ruler to measure the length of a line and then the teacher asked if they could use other tools to measure length. One student proposed to use erasers and, at the teachers' request, demonstrated how to do it.

The teacher then separated the students into two groups and gave them markers, individual whiteboards, four strips of paper (red, purple, green, and yellow), and rulers or dominoes. One group was to measure each strip with the ruler, the other with dominoes, registering their results in their individual whiteboards, while the teacher observed their work and raised questions. They then took turns in registering their findings in a large data table on the classroom wall (see Table 2).

Table 2: The data table filled by students showing lengths of strips of paper in dominoes and in inches

	Length in dominoes	Length in inches
Strip 1: yellow	6	12
Strip 2: red	2	4
Strip 3: purple	3	6
Strip 4: green	4	8

The teacher read the values on the table and asked students to determine the number of inches in one domino. A student claimed, correctly, that the length of one domino was 2 inches. Asked by the teacher to explain her answer, the student stated: *"... because like on purple it's 3 dominoes, and it's 6 inches. So, like 3 + 3 is 6, and the* *same with like the yellow, and red, and green: 6 + 6 = 12, 2 + 2 = 4, 4 + 4 = 8*". She was, in her words, referring to the fact that adding each input to itself gives the output.

The teacher asked whether the fact that 3+3=6 meant that each domino was 3 inches. The student confirmed that she still wanted it to be 2 inches; however, she was not able to relate this fact to an equation such as 6+6=12. Because the student was not immediately able to articulate how she concluded that each domino was 2 inches long, the teacher attempted to help her by asking how many 3's there are in one 6, or how many 6's in one twelve, and so on. She asked other students to join the conversation and another student tried to justify that each domino was 2 inches by referring to doubles "*... if you subtract...six minus three then you get three and that's the double*."

Next, the teacher asked the students to apply what they have found so far to a new case: the measure, in domino length units, of a 20 inches strip of paper.

A student proposed the incorrect answer 20 (dominoes). The teacher asked him to recall how many inches make the length of a domino and, again, asked for the number of dominos that would make 20 inches. Another student wrote on his whiteboard 20 x 2 = 40 and the teacher asked him to explain his answer. His explanation was that: "[I] *have 20 times 2 equals 40 because I added the dominoes and the dominoes are doubled*."

The teacher then drew the students' attention to the data table and included the number 20 in its rightmost column, along with the other numbers that happen to be doubles of the number to their left and asks: *"... if we had an orange strip of paper and it equaled 20 inches, how many dominoes could that be? In length? Looking at our pattern, we know one domino is 2 inches, what do you think?*" Then, the student who had first thought that the answer was 40, excitedly answered: *"So, it would be ten dominos.*" He also showed that he had now written 2 x 10 = 20 in his whiteboard and explained "*Because the dominos are the double* [of one inch] *and so, 2 times 10 equals 20*." Other students agree and the teacher rephrases the relationship as "*They are all times 2*."

The fact that the teacher had refocused on the chart may well have helped the students solve the problem since they could see that the missing value was associated with a position in the left column, where the lesser member of each ordered pair was located. It may have also helped that the table headers (length in dominoes and length in inches) highlighted the properties the column entries represented.

While the teacher was expecting the students to mention they were multiplying the number of dominoes by two to get to the number of inches, or dividing the number of inches to get to the number of dominoes, she did not give away this information but used the table display to help them come up with the correct answer.

It should be noted that, in her discussion with the students, the teacher sometimes focused on pure numbers (e.g., 2) to the exclusion of extensive quantities ("two inches), intensive quantities ("two inches per domino"), and statements expressing the relationship between the various quantities ("ten dominoes times two inches per domino equals twenty inches"). For students to gain an appreciation of the bearing of an equation such as $2 \times 10 = 20$ on the discussion at hand, one would hope to draw attention to relations among quantities and their units of measure[5]. We would have been happier to see her helping her students get to an answer of the sort "2 inches/domino" or "2 inches per domino length" rather than a plain 2 answer.

---

[5] Such issues were heavily emphasized in the teacher development program. However, the shift from stressing numeric computations to describing relations between quantities is admittedly a slow one.

Even though the final project lesson could have been better, the teacher made notable progress since her work with children in Course 1. At the end of the program, she elicited students' ideas and individual spontaneous representations, engaged them in discussions, and used questioning about data in a table to promote generalizations regarding any pair of input/output values.

The following is the teacher's evaluation of her own progress throughout the program, as reported in her Final Project report.

> One of the biggest "takeaways" I have from Poincaré [the program's name] is to challenge students. I always thought I had high expectations for students, however this course has taught me how to pull the information out of the students in a challenging, yet nurturing way. Throughout the three semesters I have seen students (as well as myself) learn how to struggle and have productivity through that struggle. In the past I don't believe I allowed the process to completely happen, jumping in to rescue students too quickly. It's not the number of questions we ask our students, but the value of the question we ask.
>
> In this particular scenario, I would venture to take these students and follow up with a similar activity using 3 inch dominoes and then maybe even the 2 1⁄2 inch dominoes. Although this activity produced answers that I was not expecting, I could see that all students were willing to work and not afraid of taking a risk to answer. When answers were not exactly what was planned for, using it as a teachable moment is beneficial for all.

Summary of changes in the examples of elementary school activities

At the start of the program, even with examples of how to implement the activity, this elementary school teacher did not elicit students ideas and representations, did not promote generalizations and discussions among students, and did not give students the opportunity to explore the relationships between variables (question *a*). Her use of a table to represent the implicit function in the activity was a mere exercise in computation. There were no discussions among the students and no attention to student reasoning (question *b*). During the activity, students were not given the opportunity to discuss variables and their multiple representations (question *c*).

At the end of the program we found a very different implementation of a lesson by the same elementary school teacher. She elicited students' ideas and representations, engaged them in discussion, and used data tables and questions to help students use the appropriate operations to answer a question about the relationships between variables (questions *a* and *b*). She used the table representation of the function as a pathway for students to present and discuss their ideas, occasionally correcting them. In the process, students expressed, in their own words, a general rule for the function (question *c*), thus revealing understanding of the relationship between variables, even though they did not use algebraic representations.

The middle school activity at the start of the program

The 8th grade teacher started with a description of the Candy Boxes situation and, instead of asking students to represent the problem in writing, proposed that they determine the number of candies for each protagonist: "*I want you to write down your idea as you try to figure out how many candies each of them [John and Mary] have.*" Students appeared confused with this question and one of them commented: "*All I have is a whole bunch of x's*."

Different from what was recommended in the instructions, the teacher proceeded by asking students to find a specific value for the unknown amount, *x*: "*How would you try to figure out the value of x? What does x represent?*" and "*How would you represent how many candies Mary has and how many candies John has?*"

Some students drew two boxes, labelling one as John's, the other as Mary's, with three candies on top of Mary's box. Without discussing the representations proposed by the students, the teacher reproduced one of the students' drawings and asked: "*How would you represent how many candies Mary has and how many candies John has?"* and "*What would you have to do if you don't know how many candies are in the box what would you need?*" One of the students then answered: "*One of those things. You know, whatever they're called, an x*." The teacher responds: "*A variable?*" The student agrees and proposes that Mary has x candies. The teacher then asked "*Did you have something else for Mary or did you just have x?*" A student proposed to write x for John and $x + 3$, or $3 + x$, for Mary.

After the students had proposed algebraic expressions for John's and for Mary's amounts, the teacher asked them to fill in a data table for different possible amounts. Therefore, the students just practiced a few computations and did not have the opportunity to explore data or make generalizations.

After students filled in the table with possible values for John and for Mary, the teacher returns to the algebraic representation and asks: "*How would we represent the total amount of candies John and Mary have together?*" A student proposes $x+3+x$. Others agree and the teacher asks him to simplify the expression.

A student asked if he could "… *write x times x plus 3*?" The teacher then probed the students to further think about this proposal by asking "*How would we represent something multiplied by itself?"* A student responds "*With those things on top, the exponents?*" The teacher then asked them to check if, for a few possible values in the box, the total number of candies for $x+x+3$ is the same as x^2+3. This helped students adopt the proper representation ($2x+3$).

Towards the end of the activity, the teacher further asked: "*So, in the equation, do you know the total number of candies that equals 2x+3?*" One student appropriately answered that it depended on how many candies are in each box. After further discussion they agree that one can use T for the total amount of candies and write $T = 2x+3$.

The middle school activity at the end of the program

The group the middle schoolteachers worked with adapted their Final Project lesson from a plan found in the early algebra internet site, where the students were to decide which of two phone plans was better: one that would cost 10 cents per minute for calls, the other that would cost 60 cents per month plus 5 cents per minute of talk. She chose to represent the variables as x and y, to prevent students from using m as standing for money or minutes and t as standing for time or total. To explore her students' understanding of decimals, she asked them to write and compute prices in dollars instead of cents, so that five cents was represented as 0.05, instead of 5.

During the lesson, she engaged the six students in the classroom in comparing the two functions described verbally in the problem statement, by filling out two data tables, one for each plan. The number of minutes (x) had been entered in the first column of each table, going from 0, to 2, 4, etc., up to 16, and the students were to calculate the total cost (y) for each number of minutes under each plan.

As students filled in the tables, the teacher noticed that many of them represented forty cents as 40. This gave room to the discussion transcribed below, on how to correctly represent the dollar and cents amounts in terms of decimals.

> Teacher: Forty dollars or forty cents?
> Student: Cents, cents, cents
> Teacher: Forty cents? So how do you show that in your table, because I see just forties there. Are you doing it in dollars or are you doing it in cents?
> […]

Teacher: If you are doing in dollars then you should have it as what? Should it be just 40 or should there be a decimal point?
Students at the same time: Ohhhh.
Student: Decimal point. I put the cent sign. I had it both ways. Whoops.
Teacher: So if we're doing it in dollars, because we need to stay consistent with it, then so, if we're doing it in dollars then what it should be for four minutes? What is it going to be?
Student: Point forty.

The teacher used the students' activity in producing data tables representing the written statements in the problem to help them reflect upon the decimal representation of values. After the tables were completed, the teacher engaged the students in a discussion that led them to generate the algebraic representation for each function in the problem. This was followed by the students' production of the graphs for each function in the plane. The teacher finally asked the students to decide which plan they thought would cost less if they were to speak for different number of minutes and which representation would they prefer to use to answer her questions. One student said that he would prefer the graph. The teacher further questioned the class about how the graph shows which plans would cost less, the same, or more, for different number of minutes. The students discussed how features of the graphical representation of the functions could inform the choice of phone plans, depending on number of minutes used.

Summary of changes in the examples of middle school activities

At the start of the program, even though trying to elicit students' ideas and representations, this middle school teacher posed unclear questions to the students, giving emphasis to finding a specific amount for the candies in the box and to use algebra notation. She did not promote generalizations and discussions about the relationship between the variables. Instead, once a student proposed to use algebraic notation, she focused on representing the relationship algebraically and on producing equations for the total, rather than using a variable to consider a function (questions *a* and *b*). On the positive side, she engaged students in generating an algebraic representation for the function (question *c*) and, when a wrong suggestion came about, she asked them to compute the total number of candies, using each of two representations of the implicit function, and to compare the results to data already in the table of possible values, so they could choose the correct representation.

In contrast to her work at the start of the program, the teacher's final project and classroom work reveals that she understood the main ideas promoted by the program, presenting a better plan and clear instructions. In the classroom she elicited students' ideas and paid attention to students' answers. She raised open-ended questions about data in tables that helped students correct wrong ideas regarding the decimal representation of numbers. She also used the tables as a pathway for students to produce algebraic notations and graphs for the two functions. She asked the students to explain how the different representations related to each other and to the problem situation (questions *a* and *b*). In doing so, the teacher led the students to use data tables and variables to express generalizations about the relationships between variables and to discuss how the multiple representations (verbal expressions, data tables, algebra notation, and the graphs for the two functions) were interrelated (question *c*).

Teachers' reflections on the program's contribution:

In anonymous surveys collected at the end of courses 1 and 2, 41% and 42% of the teachers' who answered the question "In what ways has this course made you think differently or more deeply?" referred to the program's contribution to how they changed their teaching by considering functions, raising open-ended questions, considering students' reasoning, and allowing students' time for reflection and understanding. Other teachers chose to comment on the courses' contribution to addressing their own difficulties and learning of mathematics.

It is worth noting that teachers were very specific in their comments rather than providing blanket statements about becoming better teachers. The following are examples of teachers' answers that addressed the courses contribution to their ways of teaching.

At the end of Course 1:

> One way that the class has changed my thinking is that every time I lesson plan my lessons I try to organize opportunities for my students to draw connections. I also respond differently to student's answers. Most of time, I ask my students why they say something or to explain how they got their answer. I am doing better with working off of student's understanding rather than pushing them in a specific direction right away. I am still trying to find a way to manage this practice with 28 students. I definitely see functions in a new way. I imagine that if we taught functions as Poincaré suggested students would have a deeper understanding of them and see their roles across math content instead of understanding them only in the confines of a unit titled functions. The course motivates me to deepen my own understandings of the content I teach. Moreover, if I had more time and the resources, I would want to read research that describe how students think about particular math ideas. It provides me with more specific guidance as to how I should support students understanding along the way.
>
> The most valuable lesson was the one on questioning (including the reading). Probing a student's understanding through questioning will undoubtedly help all teachers who implement it.
>
> This course has helped me to become a better question asker. For instance, during my classes I try to ask questions so that my students are providing their thoughts rather than providing my understanding of a topic. This has allowed students to develop their own understandings in a way that makes the most sense to them.
> I have more empathy for struggling students. As soon as they are about to have a breakthrough with understanding, the concept and content change.

At the end of Course 2:

> The connections from one topic to the next in the form of multiple representations have always been the most enlightening. It has also reinforced the need for math students to be able to struggle through difficult problems (without going too far). Support and communication are such a big part of a course like this or, for that matter, any class in math. Providing the same opportunities and support for our students is necessary to achieve the same growth that we've experienced.
>
> This course has allowed me to broaden my ways of teaching to include and allow students to struggle to create a higher level of thinking.
>
> The classroom projects, including the final created a great atmosphere in my classroom where discussing topics was easy, involvement was key. It opened my eyes to the importance of prior knowledge to help kids understand new concepts deeply.
>
> This course has continued to strengthen my teaching abilities. It continues to teach me to let the students do more of the thinking and have me take a step back

In summary

The classroom examples and teachers' survey answers suggest that changes in teaching by the end of the program match the program's focus on discussions, response to students' ideas, and a broader conceptual approach, including the use of variables and of multiple representations while teaching and discussing curriculum topics.

We next address the student assessment results.

On student assessment performance (questions d and e)

To answer research questions *d* and *e*, we examined student learning over the year teachers were taking the first two courses in the program and developing the program's activities in their classrooms. Accordingly, we analyzed PARCC's standard assessment results for academic years 2014-2015 (labeled 2015) and 2015-2016 (labeled 2016), reported by the Massachusetts Department of Elementary and Secondary Education (DOE) (http://profiles.doe.mass.edu/state_report/parcc.aspx). PARCC's results are published as the percentage of students scoring at each of five levels of achievement: L1 – Did Not Yet Meet Expectations, L2 – Partially Met Expectations, L3 – Approached Expectations, L4 – Met Expectations, and L5 – Exceeded Expectations.

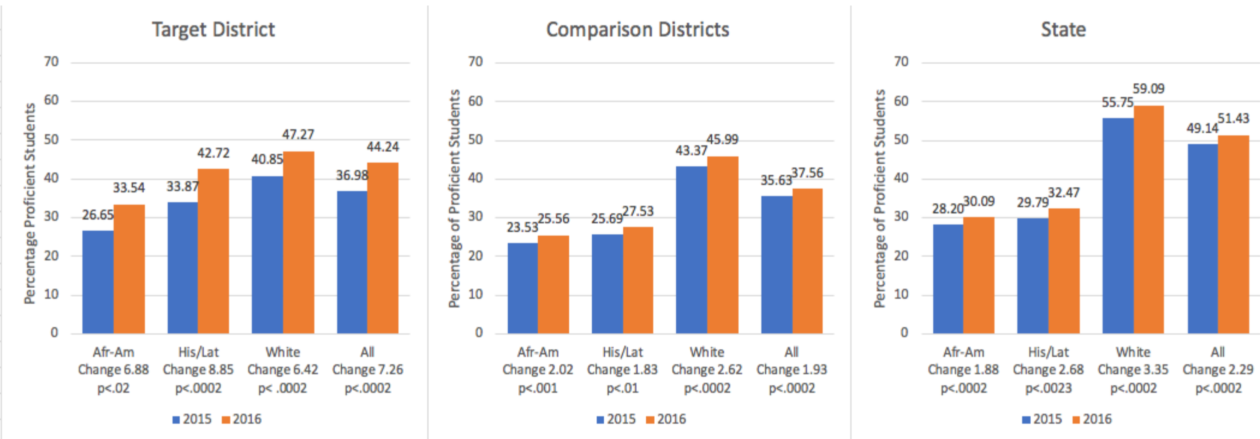
We need to acknowledge two possible extraneous variables in our analysis: (1) the students are not the same students (we compared, for example, grade 6 students in one year with grade 6 students in the following year; (2) the assessments were not necessarily comparable because there are changes from year to year. However, the law of large numbers makes interference of (1) unlikely; and the fact that we have comparisons across the state and in similar districts would seem to control for (2).

We first considered changes in the percentages of grades 3-8, for all students and for African-American, Hispanic/Latino, and White students, at the two higher proficiency levels, L4 and L5, the main index used in state reports. In addition, to examine changes across the five levels, we performed an analysis of mean levels of achievement for the same student groups. In both analyses, we compared target district results to those from the 10 comparison districts and from the whole state.

Changes in Proficiency Levels

Figure 2 shows, for each group, the percentage of students assessed as proficient (levels L4, L5) in 2015 and in 2016. The percentages in 2016 were significantly higher than those in 2015, for all students and for each ethnic groups in the target, comparison, and state districts. Significance levels for differences appear under each column pair.

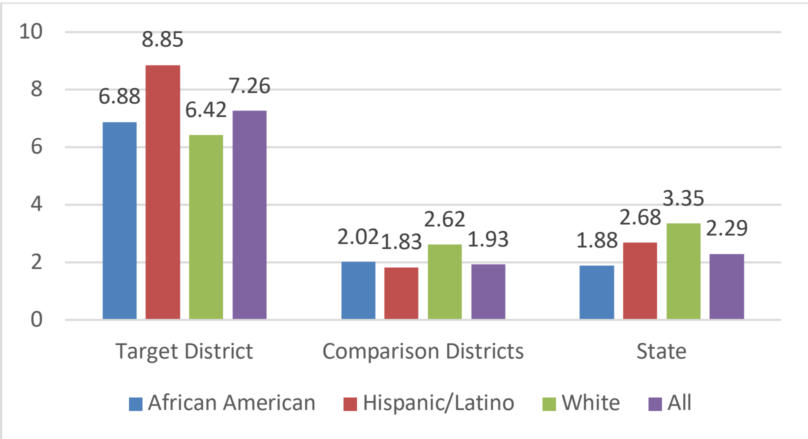
Figure 2: Percentage of proficient students in 2015 and 2016



Note: Significance was determined on the basis of z-values for differences.

Figure 3, on the relative change in the percentages of proficient students, displays the year-over-year changes.

Figure 3: Changes in the percentages of proficient students, from 2015 to 2016



The increase in the percentage of all students deemed proficient in the target district (7.26%) was over three times the increase in the comparison districts (1.93%) and in the state (2.29%). The increase among African-American students in the target district (6.88%) was more than threefold the increase for African-American students in the comparison (2.02%) and in the state districts (1.88). The increase among Hispanic/Latino students in the target district (8.85%) was more than four times the increase for Hispanic/Latino students in the comparison districts (1.83%) and more than three times the increase for the state (2.68%).

The increases in the percentages of proficient African-American and Hispanic/Latino students in the target district were larger than increases for White students. In the comparison districts and across the state, the increase in the percentage of proficient students was always larger for White students.

The achievement difference between African-American and Whites decreased slightly, by 0.47%, and the difference between Hispanic/Latinos and Whites decreased by 2.43%.

Changes in Mean Performance Level

In our second analysis, from the number of students at each level and the total number of students in each group, we computed the average performance level, treating each PARCC level as a score from 1 to 5. A group's mean performance is simply the average of its students' level scores. For example, in 2016, among the 492 African American students in the target district, 51 students performed at level L1, 117 at L2, 159 at L3, 152 at L4, and 13 at L5. The average performance for this group is, therefore, 2.917 obtained through the following calculation: $(1*51 + 2*117 + 3*159 + 4*152 + 5*13) / (51 + 117 + 159 + 152 + 13) = 1435/492$.

Table 3 displays the mean performances for each group in 2015 and in 2016, as well as changes in means from one year to the next.

Table 3. Mean performance level and changes from 2015 to 2016

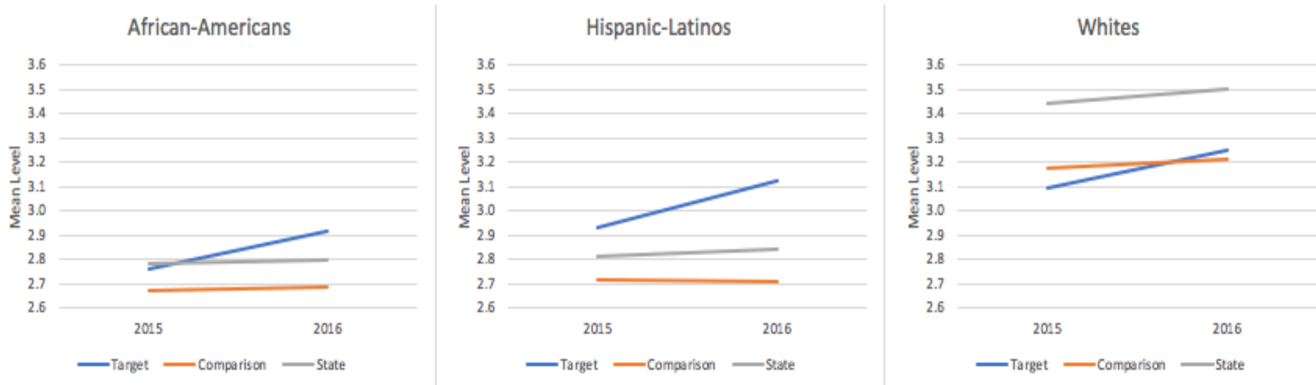
Ethnic Group	Target District 2015	2016	Change	p and ES	Comparison Districts 2015	2016	Change	p and ES	State 2015	2016	Change	p and ES
Afr-Am	2.759	2.917	0.158	<.02	2.675	2.690	0.014	n.s.	2.780	2.796	0.016	< .05
N	469	492		0.154	6,484	6,742		0.014	21,826	24,051		0.014
%	17.3	17.7			19.5	19.5			10.1	10.7		
His/Lat	2.930	3.123	0.193	<.0001	2.719	2.713	-0.006	n.s.	2.815	2.840	0.026	< .001.
N	1,116	1,215		0.189	8,130	8,867		-0.006	37,860	43,820		0.024
%	41.2	43.9			24.4	25.7			17.4	19.4		
White	3.092	3.249	0.157	<.001	3.174	3.212	0.038	<.001	3.442	3.505	0.063	<.0001
N	891	844		0.154	15,946	16,098		0.037	133109	132111		0.064
%	32.9	30.5			47.8	46.6			61.3	58.6		
Other	---	---			---	---			---	---		
N	231	156			2773	3997			24,335	25,597		
%	8.5	5.8			8.32	11.6			11.2	11.3		
All	3.009	3.165	0.156	<.0001	2.973	2.989	0.016	<.006	3.285	3.319	0.034	<.0001
N	2,707	2,769		0.152	33,333	34,557		0.015	217,130	225,579		0.032
%	100	100			100	100			100	100		

Note: Probability levels were determined by the Mann-Whitney's U test.

Consistent with findings regarding percentage of students at proficiency levels, the changes in mean performance for all students in the target district (0.156 levels) were greater than those in the comparison (0.015 levels) and in the state (0.032 levels) districts. The three ethnic groups in the target district and in the state showed significant positive changes from 2015 to 2016 for all ethnic groups; in the comparison districts, only White students showed significant improvement. Most importantly, in the target district, the change in mean performance level for African-Americans (0.158) matched the change for Whites 0.157), and the change for Hispanic/Latinos (0.193) was greater than the change for White students. In the comparison and state districts, White students showed larger changes (0.038 and 0.063 levels, respectively) than did African-American (0.014 and 0.016) and Hispanic/Latino students (-0.006 and 0.026).

An Analysis of Variance showed that the differences between mean performance levels for each of the three main factors (time of assessment, district, and ethnic group), were all significant. The specific interaction between district (target vs. comparison) and time (2015 vs. 2016) was also significant ($F_{(1, 1)} = 28.96$, $p < 0.0001$), as illustrated by the steeper slopes for the target district graphs for each ethnic group in Figure 4.

Figure 4. Mean performance level in 2015 and 2016 for each group in the target, comparison, and state districts



The meaning of our findings regarding the mean performance level

African-American students in grades 3 to 8 in the Target district had a PARCC mean performance level score of 2.759, a score that falls between levels 2 ("partially met expectations") and 3 ("approached expectations"). In 2016, the African-American average score was slightly higher, 2.917. This improvement corresponds to a difference of 0.158 levels (2.917 – 2.759), representing an average gain of, approximately, 1/7 of a PARCC level for the African-American students in the target district. Note that this gain cannot be attributed simply to the learning one expects to occur between one grade and the next because PARCC assessments and respective levels are established independently for each school grade. All things being equal, the average African-American student performance in 2016 would be expected to match the average student performance in 2015. We infer from the significance level ($p < 0.02$) that this was not the case.

Hispanic/Latino students in grades 3-8 in the Target district had a mean performance level of 2.930 in 2015, which falls within level 2 (“partially met expectations”), just shy of level 3. In 2016, their mean was 3.123, that is, just within level 3 (“approached expectations”). This improvement corresponds to a difference of 0.193 levels between the means (3.123 – 2.930). This highly significant. ($p < 0.0001$) average gain among the Hispanic-Latino students of the target district corresponds approximately to one-fifth of a PARCC level.

White students in the Target district posted mean performance gains of 0.157 (3.249 – 3.092). This difference between means was significant ($p < 0.001$).

The effect sizes in the target district (z-scores calculated by dividing the change in mean performance by the standard deviation obtained by pooling the 2015 and 2016 data) were 0.154 for African-Americans, 0.189 for Hispanic/Latinos, and 0.154 for Whites. The mean gain for all students in the Target district was highly significant ($p < 0.0001$), with an effect size (0.152) of a magnitude consistent with the effect sizes for its ethnic groups. These effect sizes, although modest, were significant. By some standards (see e.g. Cohen, 1969, p. 23), the effect size of 0.154 would be regarded as small. But as Glass et. al. (1981) have noted, even an effect size of 0.1 could be substantial when dealing with achievement gains of large numbers of students, as is the present case. Note also that these data provide conservative estimates of the impact of the program, given that only 60% of the educators from the target district were enrolled in its courses.

In the Comparison districts, neither the African-American nor the Hispanic-Latino students presented significant gains. The White students posted a significant ($p < 0.001$) but small (0.038) gain.

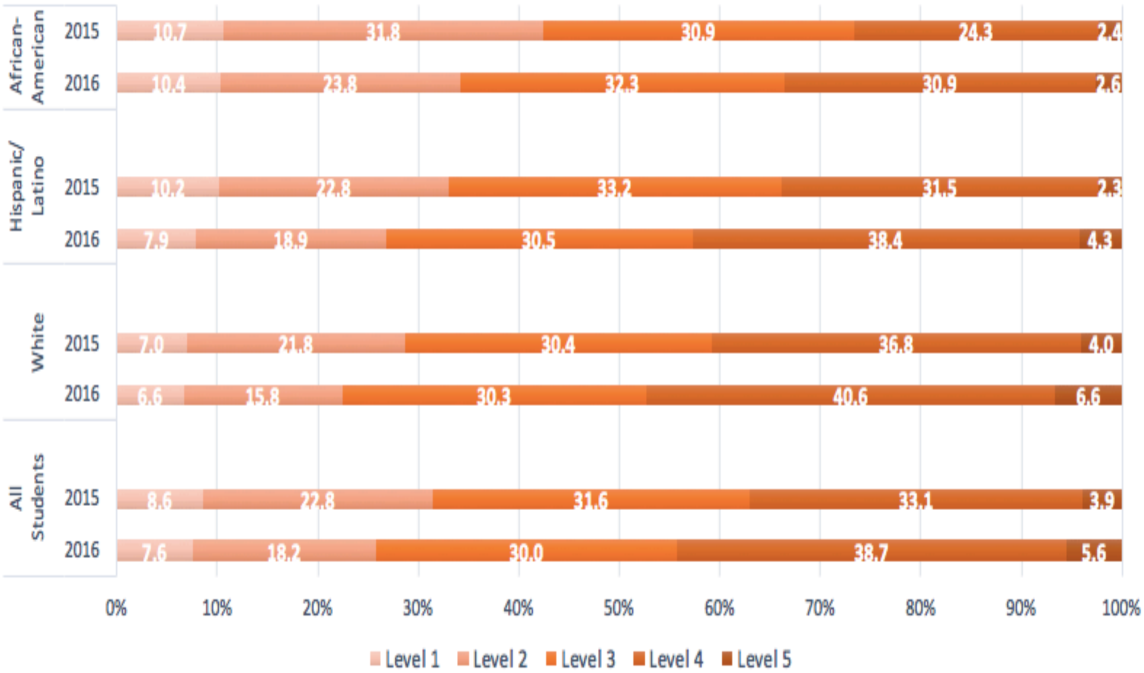
Throughout the state, the African-American and the Hispanic-Latino students presented significant but also very small gains (0.016 and 0.026 levels, respectively). The White students posted a highly significant ($p < 0.0001$) but relatively modest gain of 0.063 of a PARCC level. The gain for all students across the state was a mere 0.038 levels.

We also analyzed the data by grade, rather than as an aggregate. With the exception of African-Americans in grades 3 and 7 and White students in grade 3, for all target district's grade levels all groups showed larger increases (or smaller decreases) than was the case in the comparison groups.

Shifts Across PARCC levels

Figure 5 shows the percentage of Target district students for each ethnic group, at each PARCC level, in 2015 and in 2016. The figure highlights, from 2015 to 2016, drops in the percentages of students of all groups at levels L1 and L2 and increases in the percentages at levels L4 and L5. African-American students shifted mainly from level 2 to level 4, Hispanic/Latino students from levels 1 and 2 to levels 4 and 5, and White students from level 2 to levels 4 and 5.

Figure 5: Percentage of target district students at each PARCC's performance level in 2015 and 2016.



Changes in achievement differences: the target versus state districts

It is reasonable to compare the African-American and Hispanic/Latino students' progress in the Target district to the progress of White students in the same Target district. However, given that the White students in the Target district also benefited from the intervention, it would be more instructive to compare the results of the two ethnic groups in the Target district to the results of White students in the whole State.

In 2015 (see Table 3), the difference between Target district African-American and State White students was 0.683 (3.442 – 2.759). In 2016, the difference between the same two groups' means was 0.588 (3.505 – 2.917), thus reduced by 0.095 (0.683 – 0.588) levels. This corresponds to a reduction of achievement differences between Target African-American and State White students by 100*0.095/0.683, that is, 13.91%, or approximately one-seventh of the original difference.

In 2015 (see Figure 3) the difference between Target Hispanic-Latino and State White students' means was 0.512 (3.442 – 2.930). The final difference between Hispanic-Latino and White students was 0.382 (3.505 – 3.123). The difference was thus reduced by 0.13 levels (0.512 – 0.382). This corresponds to a reduction of the difference between Target Hispanic-Latino and State White students by 100 * 0.13/0.512, that is, 25.39%, or one-fourth of the original difference.

Summary of student learning results

For each of our analysis, gains from 2015 to 2016 for the target district's three ethnic groups outpaced those of the comparison districts and of the state. In the target district, statistically significant gains emerged for African-American, Hispanic/Latino, and White groups. Within the Comparison districts and across the state, gains were minor, with White groups showing larger gains than the minority groups.

The percentages of African-American and of Hispanic/Latino students at levels 4 and 5 showed practically no change in the achievement difference for African-Americans and a reduction of the achievement difference for Hispanic/Latino students, in comparison to White students. In contrast, in the Comparison Districts and in the State, increases in the percentage of students at levels 4 and 5 were greater for White students, resulting in increases in achievement differences between this group and the African-American and Hispanic/Latino students. Similar trends emerged regarding gains in mean performance levels.

Changes in the percentage of students across the five levels show African-American students mostly moving from level 2 to level 4 and the Hispanic/Latino and White students moving from the lower levels to levels 4 and 5.

Sixty percent of the educators in the target district participated in the program. Had all of them enrolled, the benefits of the program might have been more substantial.

Discussion

The foregoing data allow for some provisional conclusions about the impact of the program on teaching and learning and to speculate on possible contributing features of the program. However, given that the present, quasi-experimental, study did not employ random assignment of participants to treatment and control groups, any claims regarding causality must be met with a healthy degree of skepticism[6].

A major finding emerges from the relative gains in performance across the three ethnic groups in the treatment districts vis-à-vis those of comparison districts and across the state. Here, three questions arise. The first question is "Are the relative gains due to the impact of the teacher development program?" If the answer to this question is affirmative, one is led to ask, "What aspects of the program and what mechanisms were likely to be responsible for the gains?" Finally, we turn to the somewhat unusual gains by African-American and Latino/Hispanic students in the target district. Taking into account that White students in the target district showed gains over their peers in comparable districts and across the state, our third question is: "How is it that gains by

---

[6] To be sure, experiments themselves remain vulnerable to extraneous variables. Random assignment for example, does not guarantee that the treatment subjects were not more predisposed to making gains in performance than control subjects; however, it does reduce the likelihood and thereby serves as a useful control against such an intrusion. Similarly, validity issues may arise if evaluators or observers make assessments without being blind to the experimental conditions. These can arise in experiments as well as non-experimental investigations.

African-American and Latino/Hispanic students in the target district came to match or surpass those of White students in the same district?"

The gains in performance of the students from all ethnic groups in the treatment districts, vis-à-vis those of comparison districts, suggest that the program contributed to student learning as measured by standard assessment results.

Determining the factor(s) responsible for students' performance gains however, is fraught with difficulty, given the numerous features of the program. We will nonetheless attempt to speculate on this matter, focusing on prominent characteristics of the teacher development program and the described examples of changes in teaching and of teachers comments on the contribution of the program.

The program aimed to foster teachers' mathematical expertise by imparting an integrated view of the mathematics that uses functions to interconnect otherwise isolated topics and to place them in an algebraic context. It aimed to foster teachers' pedagogical expertise by familiarizing them with vivid examples of students being taught under the guidance of this view and by having them implement (and often design) classroom activities consistent with the approach.

There are indications that, as teachers became familiar with new ways of introducing the curriculum topics, they were able to explore the mathematics content in more depth. As they became more aware of a variety of approaches and representations afforded by the focus on variables, functions, and relations between quantities, they became more inclined to take students' reasoning into account even when it deviated from approaches being introduced in the textbooks. This led to students' greater participation in classroom discussions. These changes may have contributed to students' further learning and deeper understanding.

Teachers' confidence in their students' achievements may have grown after the teachers analyzed research videos of diverse classroom showing third and fourth grade students making sound generalizations and using variable notation and graphs of linear functions to solve problems. For example, two teachers expressed, in the end of Course 1 anonymous survey, that the course had:

> … challenged me to push beyond the standards to see just how far a class can go. I have been pleasantly surprised to see younger students push themselves into thinking algebraically.
>
> … made me realize how algebra needs to be presented to students at a young age.

As Rosenthal and Jacobson (1968) found, "when teachers expected that certain children would show greater intellectual development, those children did show greater intellectual development (p. 20)". This aspect may also have contributed to better learning among African-American and Hispanic/Latino students in the target district.

Institutional support may have been another factor behind the program's contribution. When invited to participate in the program, the administrators and mathematics coordinators in the district had just implemented a new curriculum to address the CCSSI requirements and adopted new textbooks. They encouraged all teachers in grades 3 to 10 to join the program. According to publicly available information from the state department of education and from districts, the comparison districts also implemented changes to address CCSSI requirements and promote their students' success. However, they did not undertake development programs as long and intensive as the one offered to the target district. The much smaller gains of comparison district students support the claim that gains in the target district are, at least in part, due to the institutional support to a long-term program focused on mathematical and pedagogical foundations.

Our analysis covered a period when the program's goals were being implemented, with (a) district support for teachers' face-to-face weekly meetings with their peers and monthly meetings with instructors, (b) teachers' online and face-to-face discussions on mathematics and on teaching, and (c) teachers' development, implementation, and analysis of classroom lessons. One question that may be asked concerns the sustainability of the changes in teaching and in student learning we have found. This aspect remains to be evaluated and possible results should guide future initiatives aimed at permanent contribution from teacher development programs.

Finally, regarding the question on why students from ethnic minorities benefitted more than white students, one possible explanation is that students from disadvantaged backgrounds, who would feel neglected in more traditional classrooms, may have been more engaged when the focus of the class moved to their own ideas and representations. Increasing teacher's mathematical knowledge and their capacity for giving all students a voice and building on their ideas created a more equitable classroom, which in turn, we speculate, were key contributors to achievement by all groups in the target district. These assumptions are consistent with results from previous studies (e.g., Boaler & Staples, 2008 and Duncan & Murnane, 2014). We further hypothesize that, as African-American and Hispanic/Latino students' mathematical ideas were valued and used as a basis for new learning, their attitudes and beliefs about their own capacity as mathematical learners may have increased.

While our results suggest that our teacher development program in mathematics may lead to more equitable teaching and narrowing of achievement differences, we must acknowledge certain limitations of our analysis:

(a) Teachers were volunteers rather than randomly selected.

(b) Results for student learning were based on state-mandated standard assessments (used before by, for example, Riordan and Noyce, 2001), which may not fully capture conceptual understanding.

(c) For confidentiality reasons, we did not have access to individual student results and were not able to separate results of students whose teachers had enrolled in the program from those of the remaining students.

(d) We did not keep track of teacher development programs and initiatives in the Comparison districts during the period we evaluated student assessment results.

(e) The effect size of gains in the target district groups, even though statistically significant, were somewhat modest.

(f) It is not clear why the Hispanic/Latino students gains were larger than those of African-Americans. Possible factors to explore in the future are the greater concentration of Hispanic/Latino students in the target district (which may make them feel at home in their school), the larger number of Hispanic/Latino teachers than African-American teachers, and the initial higher performance level of Hispanic/Latinos.

(g) We did not systematically quantify and evaluate, for the cohort of teachers in this study, changes in their classroom teaching which would help explain specific changes in performance.

(h) We did not evaluate the sustainability of the achievements we witnessed.

The program aimed at promoting better teaching and learning for all students, by adopting a somewhat novel approach to the mathematics curriculum content and by stressing pedagogical features recognized as important in addressing achievement differences. Future research should evaluate whether a program such as this, implemented at the district level over many years, with students benefiting from better teachers across several grades, would ultimately increase in a substantial way the percentage of minority students at levels 4 and 5.

The program activities did not include materials regarding ethnic group differences and were not specifically designed to close ethnic gaps in achievement. It would be useful to include such materials in future implementations of the program.

NCTM and CCSSI standards acknowledge the importance of algebra and functions throughout the curriculum. The NCTM standards emphasize key connections among variables, functions, and relations throughout the curriculum and recommends that they be explored even before middle school. These recommendations are still to become a reality in our classrooms. The effort described in this study supports the recommendations and constitute a step towards the preparation of teachers to address NCTM and CCSSI standards.

The present results point to the importance of long and intensive teacher development programs that integrates mathematical and pedagogical knowledge through the lens of functions. We are motivated to undertake studies of the program's impact over multiple years, with randomly-assigned experimental and control groups and access to individual teacher and student data. A new version of the teacher development program should also include activities aimed at raising teachers' awareness of factors behind the persistent achievement difference in our schools. So far, the program has been offered only to in-service teachers. Incorporating our approach in the pre-service teacher preparation curriculum appears as a relevant future goal.

References

Akiba, M. & Liang, G. (2016). Effects of teacher professional learning activities on student achievement growth. *The Journal of Educational Research, 109*(1), 99–110.

Akiba, M., LeTendre, G. K., & Scribner, J. P. (2007). Teacher quality, opportunity gap, and national achievement in 46 countries. *Educational Researcher, 36*(7), 369-387.

Bautista, A. (2015) (Ed.) Teacher Professional Development: International Perspectives and Approaches. *Psychology, Society and Education, Special Issue, 7*(3).

Bautista, A., & Ortega-Ruíz, R. (2015). Teacher professional development: International perspectives and approaches. *Psychology, Society and Education, 7*(3), 240-251.

Blanton, M. Brizuela, B., Gardiner, A. M., Sawrey, K., & Newman-Owens, A. (2015). A learning trajectory in 6-year-olds' thinking about generalizing functional relationships. *Journal for Research in Mathematics Education*, *46*(5), 511- 558.

Blanton, M., Stephens, A. Knuth, E., Gardiner, A. M., Isler, I., & Kim, J. -S. (2015). The development of children's algebraic thinking: The impact of a comprehensive early algebra intervention in third grade. *Journal for Research in Mathematics Education*, *46*(1), 39-87.

Bloom, B.D., Davis, A. & Hess, R. (1965). *Compensatory Education for Cultural Deprivation*. New York: Holt, Rinehart and Winston.

Boaler, J. & Staples, M. (2008). Creating Mathematical Futures through an Equitable Teaching Approach: The Case of Railside School. *Teachers College Record 110*(3), 608–645

Bol, L., & Berry, R. Q. (2005). Secondary mathematics teachers' perceptions of the achievement gap. *The High School Journal, 88*(4), 32-45.

Briars, D. & Resnick, L. (2000). *Standards, Assessments—and What Else? The Essential Elements of Standards-Based School Improvement*. CSE Technical Report 528. National Center for Research on Evaluation, Standards, and Student Testing. Los Angeles, CA: University of California.

Brizuela, B. M., Blanton, M., Sawrey, K., Newman-Owens, A., & Murphy Gardiner, A. (2015). Children's use of variables and variable notation to represent their algebraic ideas. *Mathematical Thinking and Learning*, *17*(1), 34-63.

Brooks-Gunn, J. & Duncan, G. (1997). The effects of poverty on children. *The Future of Children, 7*(2)

Brown, R. (2012). *Educators' Perspectives on Closing the Mathematics Achievement Gap in Fifth-Grade Mathematics Classrooms*. Doctoral Dissertation, Walden University.

Carraher, D. W. & Schliemann, A. D. (2018). Cultivating Early Algebraic Reasoning. In C. Kieran (Ed.), *Teaching and learning algebraic thinking with 5- 12- year-olds*. The Global Evolution of an Emerging Field of Research and Practice. ICME-13 Monographs. Chaim, Switzerland: Springer International Publishing, pp. 107-138.

Carraher, D. W. & Schliemann, A. D. (2019). Early algebraic thinking and the US mathematics standards for grades K to 5. Infancia y Aprendizaje, 42(3), pp. 479-522.

Carraher, T.N., Carraher, D.W. & Schliemann, A.D. (1985). Mathematics in the streets and in schools. *British Journal of Developmental Psychology, 3*, 21-29.

Carraher, D. W., Schliemann, A. D., & Brizuela, B. (2005). Treating operations as functions. In D. Carraher & R. Nemirovsky (Eds.), *Monographs of the Journal for Research in Mathematics Education*, XIII, CD-Rom Only Issue.

CCSSI [Common Core State Standards Initiative] (2010). *Common core state standards for mathematics*. Retrieved from http://www.corestandards.org/wp-content/uploads/ Math_Standards1.pdf.

Chazan, D. (1999). On teachers' mathematical knowledge and student exploration: A personal story about teaching a technologically supported approach to school algebra. *International Journal of Computers for Mathematical Learning 4*: 121–149.

Chazan, D. (2000) *Beyond Formulas in Mathematics and Teaching: Dynamics of the High School Algebra Classroom.* New York: Teachers College Press.

Cohen, J. (1969) Statistical Power Analysis for the Behavioral Sciences. NY: Academic Press.

Coleman, J. S. (1966). *Equality of educational opportunity*. Retrieved May 13, 2018, from http://files.eric.ed.gov/fulltext/ED012275.pdf

Darling-Hammond, L., Hyler, M. E., Gardner, M. (2017). *Effective Teacher Professional Development*. Palo Alto, CA: Learning Policy Institute.

Desimone, L.M. & Long, D. (2010). Teacher Effects and the Achievement Gap. Teachers College Record, 112(12), 3024–3073.

Desimone, L.M., Smith, T., & Phillips, K. (2013). Linking student achievement growth to professional development participation and changes in instruction: A longitudinal study of elementary students and teachers in Title I schools. Teachers College Record, 115(5), 1–46.

Desmond-Helmann (2016). What if... | A letter from the CEO of the Bill & Melinda Gates Foundation. https://www.gatesfoundation.org/2016/ceo-letter#section4 (retrieved on 3/20/19, 12:05 PM).

Duncan, G. & Murnane, R. (2014a). Growing income inequality thretens American education. *Phi Delta Kappan, 95*(6). (citation from page 14).

Duncan, G. & Murnane, R. (2014b). Meeting the educational challenges of income inequality. *Phi Delta Kappan, 95*(7).

Duncan, G. & Murnane, R. (2014c). *Restoring opportunity: The crisis of inequality and the challenge for American education.* Cambridge, MA: Harvard Education Press.

Duncan, G., Yeung, W.J., Brooks-Gunn, J.; and Smith, J.R. (1998). Much Does Childhood Poverty Affect the Life Chances of Children? *American Sociological Review, 63*(3), 406-423.

Fernandes, A., Crespo, S. & Civil, M. (2017). *Access & Equity: Promoting High-Quality Mathematics*. Reston, VA: The National Council of Teachers of Mathematics.

Franke, M. L., Carpenter, T. P., & Levi, L. (2001). Capturing teachers' generative change: A follow-up study of professional development in mathematics. *American Educational Research Journal, 38,* 653–689.

Freudenthal, H. (1973). *Mathematics as an Educational Task.* Dordrecht, The Netherlands: Reidel.

Freudenthal, H. (1991). *Revisiting mathematics education.* Dordrecht, The Netherlands: Kluwer.

Gersten, R., Taylor, M. J., Keys, T. D., Rolfhus, E., & Newman-Gonchar, R. (2014). Summary of Research on the Effectiveness of Math Professional Development Approaches. REL 2014-010. *Regional Educational Laboratory Southeast*.

Glass, G. V., McGaw, B., and Smith, M.L. (1981) *Meta-Analysis in Social Research*. Beverly Hills, CA: Sage.

Gowers, T., Barrow-Green, J., & Leader, I. (Eds.). (2010). *The Princeton companion to mathematics*. Princeton University Press.

Gravemeijer, K. (1999) How Emergent Models May Foster the Constitution of Formal Mathematics, *Mathematical Thinking and Learning, 1*(2), 155-177, DOI: 10.1207/ s15327833mtl0102_4

Gutierrez, R. (2008). A "Gap-Gazing" Fetish in Mathematics Education? Problematizing Research on the Achievement Gap. *Journal for Research in Mathematics Education, Vol. 39*, No. 4, pp. 357- 364.

Harel, G. & Dubinsky, E. (Eds.) (1992) *The concept of function: Aspects of Epistemology and Pedagogy*. Mathematical Association of America Notes, vol. 25.

Hiebert, J. and Grouws, D.A. (2007). The effects of classroom mathematics teaching on students' learning. In F. Lester (ed.) Second Handbook of Research on Mathematics Teaching and Learning: A project of the National Council of Teachers of Mathematics. Vol I. Charlotte, NC: Information Age Publishing, pp. 371-404.

Hill, H. C. (2007). Learning in the Teaching Force. *The Future of Children, 17*(1), pp. 111-127.

Hill, H. C., & Lubienski, S. T. (2007). Teachers' mathematics knowledge for teaching and school context: A study of California teachers. *Educational Policy, 21*(5), 747-768.

Hill, H. C., Blazar, D., & Lynch, K. (2015). Resources for Teaching: Examining Personal and Institutional Predictors of High-Quality Instruction. *AERA Open, 1*(4), pp. 1-23.

Hirsh, S. (2005). Professional development and closing the achievement gap. *Theory into practice, 44*(1), 38-44.

Kane, T. J., & Staiger, D. O. (2012). Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains. Seattle, WA: Bill and Melinda Gates Foundation. Retrieved from http://www.metproject.org/down-loads/MET_Gathering_Feedback_Research_Paper.pdf.

Kaput, J. (1998). Transforming algebra from an engine of inequity to an engine of mathematica l power by "algebrafying" the K-12 curriculum. In National Council of Teachcrs of Mathematics and Mathematical Sciences Education

Board Center for Science, Mathematics and Engineering Education , National Research Council (Sponsors). *The Nature and Role of Algebra in the K-14 Curriculum* (pp. 25-26). Washington, DC: National Academies Press.

Kaput, J. J. (1995). Long-term algebra reform: Democratizing access to big ideas. In B. LaCampagne, & Kaput (Eds.), The algebra initiative colloquium (pp. 37–53). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

Kirshner D. (2001) The Structural Algebra Option Revisited. In: Sutherland R., Rojano T., Bell A., Lins R. (eds) Perspectives on School Algebra. Mathematics Education Library, vol 22. Springer, Dordrecht.

LaCampagne, C. B., Blair, W., & Kaput, J. J., (Eds.). (1995). The algebra initiative colloquium: Vol 1: Plenary and reactor papers. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement. Vol 2: Working group papers ED385437.pdf Retrieved from https://files.eric.ed.gov/fulltext/ED385436.pdf .

Ladson-Billings, G. (2000). Fighting for our lives: Preparing teachers to teach African-American students. *Journal of Teacher Education, 51*(3), 206-214.

Lobato, J. & Ellis, A.B. (2002). The Teacher's Role in Supporting Students' Connections Between Realistic Situations and Conventional Symbol Systems. *Mathematics Education Research Journal, 14*(2), *99-120.*

McMeeking, L. B., Orsi, R., and Cobb, B. (2012). Effects of a Teacher Professional Development Program on the Mathematics Achievement of Middle School Students. *Journal for Research in Mathematics Education, 43*(2), 159–181.

NCTM [National Council of Teachers of Mathematics] (2000). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics. Retrieved from https://www.nctm.org/Standards-and-Positions/Principles-and- Standards/Algebra/ .

Nunes, T., Schliemann, A.D. & Carraher, D.W. (1993). Mathematics in the Streets and in Schools. Cambridge, U.K: Cambridge University Press.

Oehrtman, M. C., Carlson, M. P., & Thompson, P. W. (2008). Foundational reasoning abilities that promote coherence in students' understandings of function. In M. P. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and practice in undergraduate mathematics* (pp. 27-42). Washington, DC: Mathematical Association of America.

Olive, J. and Caglayan, G. (2008). Learners' Difficulties with Quantitative Units in Algebraic Word Problems and the Teacher's Interpretation of those Difficulties. International Journal of Science and Mathematics Education. Vol. 6(2), pp. 269-292.

Ostashevsky, L. (2019). More than five years after adopting Common Core, Kentucky's black-white achievement gap is widening. The Hechinger Report: https://hechingerreport.org/five-years-adopting-common-core-kentuckys-black-white-achievement-gap-widening/ .

Piaget, J. (1995). *Sociological Studies*. London and New York, Routledge.

Piaget, J. & Inhelder, B. (1974). *The child's construction of quantities : conservation and atomism*. London, Routledge and Kegan Paul.

ProCon.org. (n.d.). Standardized Tests. Retrieved from https://standardizedtests.procon.org/)

Reed, H. J. & Lave, J. (1979). Arithmetic as a tool for investigating relations between culture and cognition. American Anthropologist, 6, 568-582.

Riordan, J. and Noyce, P. (2001). The Impact of Two Standards-Based Mathematics Curricula on Student Achievement in Massachusetts. Journal for Research in Mathematics Education 2001, Vol. 32, No. 4, 368–398.

Rosenthal, R. & Jacobson, L. (1968). Pygmalion in the classroom. *Urban Review, 3*(1), 16-20.

Santagata, R., Kersting, N., Givvin, K. B., & Stigler, J. W. (2011). Problem implementation as a lever for change: An experimental study of the effects of a professional development program on students' mathematics learning. *Journal of Research on Educational Effectiveness, 4*(1), 1-24.

Saxe, G. (1991). *Culture and Cognitive Development: Studies in Mathematical Understanding*. New York, Psychology Press.

Saxe, G. B., Gearhart, M., & Nasir, N. S. (2001). Enhancing students' understanding of mathematics: A study of three contrasting approaches to professional support. *Journal of Mathematics Teacher Education, 4*, 55–79.

Schliemann, A. D., Carraher, D. W., & Brizuela, B. M. (2006). *Bringing out the algebraic character of arithmetic: From children's ideas to classroom practice*. Hillsdale, NJ: Lawrence Erlbaum Associates

Schliemann A. D., Carraher D. W., Brizuela B. M. (2012). Algebra in elementary school. In L. Coulange & J.-P. Drouhard (Eds.) Enseignement de l'algèbre élémentaire: Bilan et perspectives. Special Issue of *Recherches en Didactique des Mathématiques*, pp. 109-124.

Schliemann, A.D., Carraher, D.W., & Teixidor-i-Bigas, M (2016). Teacher Development and Student Learning. Invited Presentation. 13th International Congress on Mathematical Education. Hamburg, Germany, July, 25-30.

Schoenfeld, A. (1995). Report of working group 1. In C. B. LaCampagne, W. Blair, & J. J. Kaput (Eds.), *The algebra initiative colloquium: Vol 2* ED385437.pdf#page=11 (pp. 11–18). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

Schoenfeld, A. H. (2002). Making mathematics work for all children: Issues of standards, testing, and equity. *Educational Researcher, 31*(1), 13-25.

Schwartz, J. & Yerushalmy, M. (1992). Getting students to function on and with algebra. In E. Harel & G. Dubinsky (Eds.), *The Concept of Function: Aspects of Epistemology and Pedagogy* (pp. 261-289). Washington, DC: Mathematical Association of America.

Seldon, A., & Seldon, J. (1992). Research perspectives on conceptions of function: Summary and overview. In E. Dubinsky & G. Harel (Eds.), *Concept of function: Aspects of epistemology and pedagogy* (pp. 1–21). Washington, DC: Mathematical Association of America

Singham, M. (2003). The achievement gap: Myths and reality. *Phi Delta Kappan, 84*(8), 586-591.

Smith, J. & Thompson, P. (2008). Quantitative reasoning and the development of algebraic reasoning. In J. Kaput. D. Carraher, & M. Blanton (Eds.), *Algebra in the Early Grades*. Mahwah, NJ, Erlbaum, pp. 95-132.

Steele, M.D. & Hillen, A.F. (2012). The Content-Focused Methods Course: A Model for Integrating Pedagogy and Mathematics Content. *Mathematics Teacher Educator*, 1(1) 53-70

Steele, M.D., Hillen, A.F. & Smith, M.S. (2013). Developing mathematical knowledge for teaching in a methods course: the case of function. *Journal of Mathematics Teacher Education, 16*, 451–482.

Sztajn, P., Borko, H., & Smith, T. (2017). Research on Mathematics Professional Development. In J. Cai (Ed.) *Compendium for Research in Mathematics Education*. Reston, VA: National Council of Teachers of Mathematics.

Teixidor-i-Bigas, M., Carraher, D. W. & Schliemann, A. D. (2013). Integrating Disciplinary Perspectives: The Poincaré Institute for Mathematics Education. *The Mathematics Enthusiast, 10*(3).

Thompson, P. W. (2011). Quantitative reasoning and mathematical modeling. In L. L. Hatfield, S. Chamberlain & S. Belbase (Eds.), *New perspectives and directions for collaborative research in mathematics education.* WISDOMe Mongraphs (Vol. 1, pp. 33- 57). Laramie, WY: University of Wyoming.

Thompson, P. W. (2015). Researching mathematical meanings for teaching. In English, L., & Kirshner, D. (Eds.), *Third Handbook of International Research in Mathematics Education* (pp. 968-1002). London: Taylor and Francis.

Vygotsky, L.S. (1978). *Mind in Society*. Cambridge, MA: Harvard University Press.