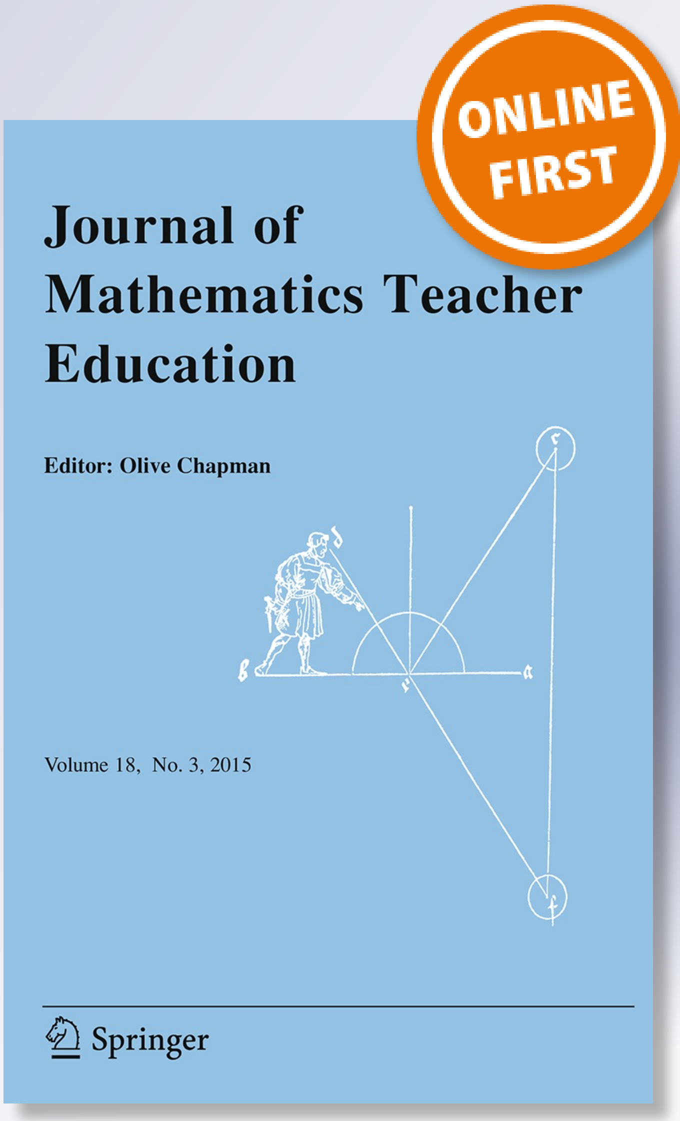
Sustainable changes in teacher practices: a longitudinal analysis of the classroom practices of high school mathematics teachers

Yasemin Copur-Gencturk & Anne Papakonstantinou

Journal of Mathematics Teacher Education

ISSN 1386-4416

J Math Teacher Educ
DOI 10.1007/s10857-015-9310-2



Springer

J Math Teacher Educ
DOI 10.1007/s10857-015-9310-2

Sustainable changes in teacher practices: a longitudinal analysis of the classroom practices of high school mathematics teachers

Yasemin Copur-Gencturk[1] · **Anne Papakonstantinou**[2]

Abstract This study examines the effects of professional development on various aspects of teachers' mathematics instruction. Using data collected between 2005 and 2009, we examined the extent to which the instructional practices of 49 US high school teachers who participated in content-based, sustained professional development changed over time. We found that changes in several aspects of their instructional practices followed somewhat different patterns. Teachers made statistically significant and steady changes in mathematical discourse, instructional clarity, and the development of students' mathematical habit of mind, but not in student interactions or in the use of multiple representations.

Keywords Professional development · In-service teacher education · High school teacher · Mathematics · Instructional practice · Longitudinal study

Introduction

Effective instruction is critical for promoting students' conceptual understanding of mathematics (e.g., Anthony and Walshaw 2009; National Council of Teachers of Mathematics [NCTM] 2000; Schoen 2003; Senk and Thompson 2003). However, even two decades after standards for teaching mathematics were introduced (e.g., NCTM 1991, 2000), high school mathematics instruction in the USA has still not created learning environments where students are actively engaged with mathematics and where the needs of diverse learners are met (e.g., Smith 2013; Weiss et al. 2003; Whittington 2002).

✉ Yasemin Copur-Gencturk
ycopur-gencturk@uh.edu

[1] Department of Curriculum and Instruction, University of Houston, 434 Farish Hall, 4800 Calhoun Rd., Houston, TX 77204, USA

[2] Rice University School Mathematics Project, Rice University, Houston, TX, USA

Published online: 21 June 2015

Teachers' movement toward reform-based mathematics teaching[1] has been limited mainly because they must have considerable knowledge of mathematics and mathematics teaching (e.g., Ball et al. 2008) and must learn many new skills to implement reform-based practices. High-quality professional development can provide the support and guidance teachers need; however, little is known about what makes professional development effective in terms of fostering teacher change. Various research agencies and national organizations have listed key aspects of effective professional development (for a review, see Guskey 2003), and these lists have shown few overlapping features. Beyond that, empirical evidence identifying consistently agreed-on features of professional development is limited (Blank and de las Alas 2010; Gersten et al. 2014; Parise and Spillane 2010; Scher and O'Reilly 2009). Therefore, further empirical evidence is needed on the features of professional learning opportunities that promote teacher change.

The purpose of this study was to contribute to the existing research by investigating how high school teachers' instructional practices improved when they participated in a program that was grounded in the most commonly agreed-on features of effective professional development, and which of these changes were sustained over time. To do so, this study addressed two shortcomings of the existing research. First, we utilized direct measures of teachers' mathematics instruction. Most studies on the role of professional development in changes in teachers' practices have been based on teachers' self-reports (e.g., Banilower et al. 2007; Desimone et al. 2013; Garet et al. 2001; Merrill et al. 2010), and discrepancies have been found between teachers' self-reported practices and those observed by others (e.g., Copur-Gencturk et al. 2014; Polly and Hannafin 2011). Second, we investigated the long-term impact of the professional development program. Only a few studies have investigated whether professional development programs have a sustained effect on teachers' practices (Copur-Gencturk et al. 2014; Boston and Smith 2011). Without knowing whether improvements in teachers' practices remain stable over time, and if so, which ones, researchers will not truly know the contribution of professional development to teacher change.

Conceptual framework

In this study, we investigated the effects of a professional development program for high school mathematics teachers on their instruction, and specifically, in relation to teachers' use of reform-oriented practices as envisioned in several of the *Standards* set forth by the NCTM (1991, 2000, 2014). With the introduction of reform-based mathematics instruction, the focus was shifted away from the traditional emphasis on the teacher as a transmitter of knowledge and from the teaching of rules to the teacher as a facilitator who assists students to construct and make sense of key mathematical concepts and procedures through active involvement in the learning process.

Reform-based teaching also places strong emphasis on mathematical discourse and the incorporation of varied mathematical representations, with specific attention being given to concrete representations in an effort to advance students' understanding and sense-making of mathematical concepts (NCTM 1991, 2000, 2014). Mathematical discourse appears to increase students' ability to use reasoning, proofs, explanations, and justification (Lampert

[1] By reform-based mathematics teaching, we are referring to inquiry-based teaching in which teachers create an environment where students actively engage in problem-solving activities and make sense of mathematical concepts through investigation and discussion.

1990; Yackel and Cobb 1996) and to increase their ability to become better problem solvers (Woodward et al. 2012).

Given that in the USA, reform-based mathematics teaching is drastically different from traditional teacher-centered instruction, in-service teachers need opportunities to relearn how to teach mathematics by using methods more closely aligned with the NCTM *Standards*. With this goal in mind, our approach to designing a professional development program was guided by general models of teacher change in professional development (e.g., Desimone 2009; Fishman et al. 2003; Guskey 2002) and by research on professional development. Our aim was to identify key features of effective professional development programs as well as key aspects of teachers' mathematical knowledge and their relation to instruction to determine which components of mathematical knowledge should be the foci of the program (e.g., Copur-Gencturk 2015; Carpenter et al. 1989).

To study the link between professional development and changes in teachers' practices, we followed Desimone's (2009) suggestion of measuring the core features of professional development activities as one way to investigate the effects of professional development programs on teachers. Thus, we identified key features of a long-term intervention program designed for high school teachers and examined the extent to which aspects of their mathematics teaching changed after they had participated in a professional development program. In the following section, we briefly summarize research on the core features of the program: its focus on content and pedagogical content knowledge, its provision of active learning experiences, its connections to teachers' work, and its duration.

Key features of high-quality professional development

Content and pedagogical content knowledge focus

One of the most commonly agreed-on features of professional development (Desimone 2009; Guskey 2003; Loucks-Horsley and Matsumoto 1999) is the focus on both content and pedagogical content knowledge. Beyond the fact that teachers cannot teach a concept if they do not know it, teachers who do not have a conceptual understanding of mathematical concepts do not seem to provide opportunities for their students to learn with understanding.

Research on teachers' mathematical knowledge and instruction indicates that compared with teachers with firm mathematical knowledge, teachers with limited mathematical knowledge, are less likely to engage in high-quality mathematics instruction (e.g., Copur-Gencturk 2015; Hill et al. 2008). As teachers' mathematical knowledge increases, they are more likely to change their lesson structure toward inquiry-based teaching, create a more welcoming classroom environment, and, more important, become more efficient in making explicit what their students are expected to learn from a lesson (Copur-Gencturk 2015). Increasing teachers' knowledge and awareness of students' thinking also appears to improve the mathematical quality of their instruction (Carpenter et al. 1989) and seems to have a sustained impact even years after the program has ended (e.g., Franke et al. 2001). Therefore, a positive change in teachers' mathematics instruction is to be expected if specific attention is given to the mathematics content and to students' thinking and learning processes.

Research also provides empirical evidence for the impact of content-focused programs on changes in the quality of mathematics instruction (e.g., Boston and Smith 2009; Farmer et al. 2003; Franke et al. 2001; Ingvarson et al. 2005; Parise and Spillane 2010; Supovitz et al. 2000). Although the majority of the professional development programs studied were designed for elementary and middle school teachers, we would expect to see a similar positive impact of content-focused professional development on high school teachers' instruction, especially given that 10 % of US high school mathematics teachers had never had a course in mathematics teaching and that 51 % had not completed a course in mathematics education in the last 10 years (Smith 2013).

Active learning experiences

There is broad consensus that professional development programs should offer opportunities for teachers to become actively engaged in a meaningful analysis of their teaching and their students' work and that the teachers should experience learning through inquiry, sharing, and discussing their ideas (Loucks-Horsley et al. 2010; Loucks-Horsley and Matsumoto 1999). Research also provides evidence for the importance of teachers being actively engaged in their own learning (e.g., Desimone et al. 2002; Farmer et al. 2003). The large-scale analysis by Ingvarson et al. (2005) of the impact of four evaluation studies with various activities on 3250 Australian teachers indicated that the active learning experiences the teachers had were among the most important influences they reported in changing their practices.

Connections to teachers' work

Several studies point to the importance of professional development in addressing the needs of participating teachers and providing learning experiences linked to their daily work of teaching (e.g., Borko et al. 1997; Boston and Smith 2009; Bullough et al. 1997; Klein 2001). Meeting teachers' needs also includes connecting professional development activities to teachers' classroom practices and their curricular contexts (Desimone 2009). Results of a correlation analysis by Garet et al. (2001) of a variety of professional development programs supported the idea that changes in teachers' practices were related to how closely the content of professional development activities was aligned with topics emphasized in the state and district standards. Similarly, Penuel et al. (2007) found that program coherence and alignment with standards had a positive impact on instruction.

Program duration

Findings from earlier research indicate that teachers often need more time and more continued support to sustain changes in their practices than what is provided through traditional, relatively short professional development opportunities (e.g., Abell and Pizzini 1992; Gibbons et al. 1997; Sandholtz and Ringstaff 2013). Sandholtz (2002) reported that teachers stressed the importance of ongoing support for sustaining changes in their practices. Similarly, research by Supovitz and Turner (2000) with science teachers indicated that teachers made the most in-depth improvements in their practices only in professional development programs of longer duration. These changes included asking students to provide evidence to support their claims and encouraging students to explain their ideas, rather than structural changes such as incorporating hands-on activities into their teaching.

Short-term programs or programs with a limited number of contact hours seemed less likely to lead teachers to improve their practices (Polly and Hannafin 2011). In a meta-analysis by Scher and O'Reilly (2009), the authors found that professional development programs in mathematics had a more profound effect on student learning if they lasted longer than one academic year. One important aspect of the length of the program was related to the support provided during that time. Guskey and Yoon (2009) argued, based on a review of nine studies, that sustained follow-up activities are key to the changes in teachers' practices and their students' learning. In longitudinal professional development programs, ongoing support throughout the academic year appears to help teachers use more open-ended questions and encourage students to consider alterative explanations (e.g., Sandholtz and Ringstaff 2013).

Limitations of existing studies

The present literature on the impact of professional development activities has three main limitations. First, the sustained effects of teachers' professional development programs have not been investigated in depth (for a review, see Avalos 2011), and only a limited number of studies have investigated the extent to which teachers have retained the changes in their practices (e.g., Boston and Smith 2011; Franke et al. 2001; Supovitz et al. 2000). Second, most existing studies have relied on teachers' self-reports to investigate the impact of professional development on their teaching (e.g., Banilower et al. 2007; Garet et al. 2001; Parise and Spillane 2010; Supovitz and Turner 2000). Studies using multiple data sources have indicated that teachers might not be able to reflect accurately on the quality of changes in their practices (e.g., Copur-Gencturk et al. 2014; Lee et al. 2004). As Guskey (2002) argued, direct observations could provide the most accurate information. However, Moyer-Packenham et al. (2011) reported, based on an analysis of more than 2000 professional development activities created by the National Science Foundation's Math and Science Partnership Programs, that the impacts of only 4 % of these activities had been examined by conducting classroom observations. Therefore, more studies utilizing direct observations are needed to better understand the impact of professional development on teachers' instructional practices.

Finally, features of effective professional development have been identified mainly based on studies of programs designed for elementary or middle school teachers (e.g., Blank and de las Alas 2010; Borko et al. 1997; McMeeking et al. 2012; Polly and Hannafin 2011; Parise and Spillane 2010; Scher and O'Reilly 2009). More research with high school teachers is necessary to examine whether the agreed-on features of effective professional development would have similar impacts on high school teachers' practices. In addition, most of the existing studies have not investigated how various features of mathematics instruction have changed, which has limited our understanding of which aspects of teachers' instructional practices would be improved through professional development.

In the present study, we aimed to address these shortcomings in the literature. We examined the extent to which a longitudinal professional development program based on the aforementioned key aspects of high-quality professional development (i.e., an extended program that focused on content and pedagogical content, that incorporated active learning experiences, and that was in alignment with current teaching and learning standards at the national, state, and district levels) was able to produce instructional changes for teachers in low-socioeconomic-status schools. We also investigated which teaching practices appeared to change immediately, which practices shifted over time, and which seemed the most resistant to change.

Methods

Project context

The program, funded by the National Science Foundation's Math and Science Partnership Program from 2004 to 2009, was designed to provide professional development for high school mathematics teachers in the USA. The professional development program was designed to prepare high school teachers—all of whom had been selected by their campus principals as lead teachers—to take the initiative in their schools to improve the quality of mathematics education. A mathematics education center, which was housed at a private university in a large southern city in the USA, designed the program in collaboration with faculty from the university's mathematics department as well as administrators from the participating school districts to addresses specific needs of teachers in the districts. The program was intended to serve as a catalyst to initiate change at the grassroots level in local schools and to influence the type and direction of mathematics instruction in the participating schools and school districts.

The center incorporated the key features of high-quality professional development. These included deepening teachers' subject-matter knowledge and pedagogical content knowledge; actively engaging teachers in inquiry-based activities; making explicit links to teachers' daily work by aligning the focus of the program with the school, district, and national reforms, policies, and guidelines; and providing professional development of sufficient duration to initiate changes in their practices (American Federation of Teachers 2002; Desimone 2009; Guskey 2002; National Academy of Education 2009).

The program was developed for three cohorts of teachers[2] over 6 years in two urban districts that mainly served low-income students or students of color. Cohorts began the program 2 years apart. The teachers in each cohort were expected to complete two consecutive summer institutes[3] designed to enhance teachers' knowledge of content and pedagogical content knowledge through inquiry-based activities. Teachers who completed each summer institute along with the following year's activities earned 4 hour of graduate credit each year. To remain active members of the program, the teachers in previous cohorts had to continue to participate in follow-up academic-year activities with subsequent cohorts.

The summer institutes across cohorts had the same format. They spanned 7 hours per day for 4 weeks each summer during two consecutive summers and were grounded in social constructivist approaches, with teachers actively engaging in collaborative group activities. The content of the summer institutes was designed by the center and included senior teachers and faculty from the mathematics department to ensure that the content addressed the big ideas in mathematics and was aligned with standards at the national, state, and district levels. Specifically, attention was given to geometry, linear algebra, and statistics and probability, given that US teachers at the secondary school level did not feel well prepared to teach some key topics such as statistics and probability (Smith 2013) and that algebra and geometry are the main foci of the high school curriculum in the USA (NCTM 2000). Therefore, during the first summer for each cohort, the mathematics content

[2] This study included only the first two cohorts because the goal of the study was to investigate the long-term effects of the program. Classroom observations conducted with the teachers in the third cohort were not longitudinal.

[3] Depending on the needs of the cohort, the content covered in summer institutes varied by cohort to some extent.

strand focused on concepts and important ideas related to geometry and linear algebra and on their connection to the curriculum in high school algebra and geometry courses. The curriculum for the algebra strand focused on understanding equivalence and the meaning of solutions, whereas the curriculum for the geometry strand was centered around a professional development module developed by the center for the state's education agency and was grounded in the work of Van Hiele and Van Hiele-Geldof (1958) and Van Hiele (1986).

The content strands for each cohort's second summer institute were combinatorics and statistics and were intended to develop teachers' understanding of permutations and combinations. After each combinatorics session, teachers explored how the problem-solving process could be integrated into their classrooms through the development of open-ended questions tied to state assessments. The statistics strand engaged participating teachers in descriptive and inferential statistics activities designed to help them see the connections between statistics and the content in high school mathematics courses. In alignment with the content focus of each summer institute, the program aimed to increase teachers' pedagogical content knowledge related to how students learn and the difficulties associated with mastery of these specific concepts, and how teachers build on what their students already know and can do.

A typical institute day involved teachers' enhancing their content and pedagogical content knowledge in the morning and addressing and exploring topics related to mathematics leadership and diversity in the afternoon. Lunch sessions provided opportunities for faculty, postdoctoral researchers, and graduate students in mathematics and statistics to share their research with the participating teachers. In the morning sessions, teachers took active roles as learners and worked collaboratively in groups. The center staff modeled the appropriate use of technology and manipulatives. After the instructional activities were completed, the underlying pedagogical approaches were deconstructed, and issues around effective student learning were examined. Particular attention was given to appropriate questioning techniques and their role in effective formative assessment.

The afternoon sessions aimed to prepare the participating teachers as leaders in their schools to initiate changes in how mathematics is taught and how students from different backgrounds can access it. The leadership component, led by the partner school districts' mathematics directors, was built on Leinwand's (2000) *Sensible Mathematics: A Guide for School Leaders* and Lambert's (1998) *Building Leadership Capacity in Schools*, as well as on case studies from Miller et al.'s (2000) *Teacher Leadership in Mathematics and Science*. Guest speakers and panels addressed underrepresentation in mathematics and science among ethnic minorities and females in undergraduate and graduate programs, professional careers, and academics. In the second summer, afternoon sessions were devoted to classroom management from a diversity perspective. These sessions aimed to provide teachers with the skills and knowledge needed to work with other teachers on classroom management issues. Issues of diversity were explored through discussions with panels of participating teachers. Topics included how to make algebra important for all students, how to motivate at-risk students, how to get students to value education, and how to deal with diverse teachers and students. In addition, each summer, groups of participating teachers studied topics of their own choice related to sociological or educational issues, or both, within their schools or as they related to mathematics, such as by addressing the achievement gap between different student groups and by exploring the relationship between music and mathematics.

Summer institutes were followed by academic-year support for the duration of the program. Support was provided both individually, in the form of campus visits, and to the group as a whole, through scheduled academic-year meetings as the program progressed, which brought together teachers from different cohorts. Each participating teacher was assigned to an educator in the center who served as a liaison. Liaisons made campus visits to mentor participating teachers and to support their campus-based instructional and leadership activities.

During the fall semester of their first year in the program, the liaisons visited each participating teacher at least once to observe the extent to which the participating teacher implemented the aforementioned reform-based teaching practices. The purpose of these visits was to identify teachers who needed further help. If the liaison felt the teacher was able to implement the reform-based practices, then no further observations were scheduled. Participating teachers who needed to adjust their practices were visited more frequently. During the second year, the focus was on each participating teacher's work as a mentor and coach for the other teachers in his or her department. In addition to visiting classes, liaisons attended collaborative planning meetings facilitated by the lead teachers. The teachers who were effectively facilitating meaningful collaborative planning were visited only once. Teachers who needed assistance in developing effective collaborative planning were visited more frequently. Each year, approximately half of the teachers were visited more than once. Two or three times each year, participating teachers would request an additional site visit to help mediate a problem or issue on their campus.

Academic-year meetings were aimed at meeting the needs of participating teachers as identified by liaisons through site visits, conversations with participating teachers, and the results of post-program and follow-up surveys. These 2-hour meetings, which were held six times each year after school hours, brought together all actively participating teachers regardless of their cohort and were facilitated by the center personnel. Each meeting had a specific focus that varied from meeting to meeting. Participating teachers identified the development of student problem-solving strategies as a primary concern; therefore, several meetings were devoted to this topic. Other meetings addressed curriculum issues, equity and diversity issues, assessment and instructional strategies, and effective communication techniques.

Participants

In the first 4 years of the program, 64 high school mathematics teachers in the first two cohorts participated (36 and 28 teachers from Cohorts 1 and 2, respectively). The program followed a similar format for each cohort. Of the 64 teachers in the two cohorts, 49 were included in this study. The remaining teachers were excluded from data collection for several reasons, such as relocation, new academic positions, and health problems.

The 49 teachers were employed in 34 high schools across the two districts. The majority of teachers were female (79 %). Almost half of the teachers were African-American (47 %), and 23 % were White (and non-Latino/Latina). All teachers were certified, and 36 held master's degrees. They taught 9th–12th grades. The teachers' years of experience in the classroom ranged from 1 to 49 years (mean = 14.1; median = 12).

The classroom observation instrument

The research team of the mathematics education center and the external evaluator of the program developed an observation instrument designed to capture the instructional changes that were expected as a result of the professional development. The instrument was grounded in the social constructivist approaches developed in the work of Dewey (1938/

1997), Piaget (Bell-Gredler 1986), Vygotsky (1962), Van Hiele (1986), and Skemp (1987), and the measured practices aligned with reform-based mathematics teaching, including providing students with opportunities to actively engage in problem-solving activities, share their ideas, and discuss and reflect on those ideas. To increase the validity and reliability of the instrument, items were tested and modified during classroom observations of teachers in previous programs developed by the center. After the initial development of the instrument, the center's research team conducted observations to verify that the items captured what they were expected to measure and to establish inter-rater reliability.

The instrument included several items intended to capture all aspects of and factors related to instruction, from capturing students' seating patterns to the types of posters placed on the walls, to the types of resources used (e.g., manipulatives, calculators, and SMART Boards), to the content focus (e.g., propositional knowledge vs. procedural knowledge), to teachers' behaviors (e.g., teachers' use of scaffolding), and to students' observed behaviors (e.g., whether students justified their conclusions). The protocol was composed of dichotomous items as well as items on a six-point frequency scale, ranging from *never* to *very often*, to identify and rate the quality and extent to which various parameters supported reform-based instruction of high school mathematics. Data analysis included 19 items that were on a six-point scale and 21 items that were on dichotomous scale.[4]

Data collection

After the teachers had completed their first summer in the program, external evaluators conducted classroom observations beginning in the fall of 2005 and continuing through the spring of 2009. The aim was to observe each teacher at least once a year. After each observation, the observed lessons were scored by using the classroom observation protocol. In total, 185 classroom observations were included in the data analysis.

We conducted factor analyses to identify the distinct features of reform-based mathematics instruction captured by our observation instrument. Given that items were on two different frequency scales (Likert and binary), two separate factor analyses were conducted. For the dichotomous data, we conducted a factor analysis of tetrachoric correlations.[5] The five scales (three from the Likert items and two from the dichotomous items) that emerged from the factor analysis were utilized in the study.[6] The names of the scales and a sample item for each are shown in Table 1. The reliabilities of the scales [Cronbach's alpha or Kuder–Richardson formula 20 (KR-20) for binary items] ranged from .74 to .87, implying acceptable consistency within each scale.

The sample items from each scale in Table 1 provide a glimpse of what each scale measured. Specifically, the Student Interactions scale captured the extent to which the teachers provided opportunities and time for students to learn from each other and whether

[4] Throughout the duration of the program, the external evaluator of the project reworded or replaced some items. However, the 40 items that stayed the same in all versions of the instruments were included in the data analysis.

[5] Given that factor analysis is dependent on correlations among items, using tetrachoric correlations is more appropriate for binary items. Therefore, we preferred to use two separate factor analyses. In both analyses, we used principal components analysis with varimax rotation and retained only factors with eigenvalues >1.

[6] Classroom climate scale measuring the climate of respect of teachers and their students for each other was not included in the data analysis because teachers' average score on the baseline measure was .9 out of a maximum of 1, which did not leave room for growth on this scale.

Table 1 Scales obtained from the classroom observation protocol

Scale	Reliability estimate	No. of items	Sample item(s)
Student Interactions	.87	5	Students discussed and explained their understandings of each question with a partner or within a small group
Mathematical Discourse	.75	4	Teacher used probing questions to deepen students' mathematical understanding
Instructional Clarity	.81	3	Teacher provided explicit expectations for group activity and product(s)
Mathematical Habit of Mind[a]	.79	6	Students were actively engaged in thought-provoking activities that often involved the critical assessment of procedures
Multiple Representations	.74	4	Students used a variety of means to represent concepts (for example, models, drawings, graphs, and manipulatives)

[a] The Kuder–Richardson formula 20 (KR-20) was estimated for the Mathematical Habit of Mind and Multiple Representations scales. The first three scales consisted of items on a six-point scale, whereas the last two included binary items

students actually used that time and those opportunities efficiently. It assessed the extent to which teachers provided opportunities and allocated class time for students to work with a partner or in a small group and to discuss their understandings with each other. Recall that student interaction is an important aspect of reform-based teaching practices in that it gives students an opportunity to exchange ideas and to learn how to engage in mathematical argumentation and validation. Thus, this scale captured the extent to which teachers changed their instructional practices to provide opportunities for students to learn from each other. The Mathematical Discourse scale captured the mathematical quality of discourse by measuring the extent to which students justified their solutions mathematically and shared their prior knowledge of the concept, as well as the extent to which their talk was about mathematics. It also assessed the extent to which the teachers used questioning strategies to deepen their students' understanding of mathematical concepts by guiding their thinking and clarifying their understanding rather than evaluating the students' solutions and providing correct answers. The Instructional Clarity scale captured the extent to which teachers made explicit to their students what they were expected to learn from the activities and the lesson. It also measured the extent to which teachers used the whole-class discussion to go over the questions, and the extent to which teachers provided explicit expectations for activities and then facilitated the whole-class discussion following the group activity to ensure that students learned what they were expected to learn from it.

Of the dichotomous scales, the Mathematical Habit of Mind scale captured the extent to which teachers created an environment where students were actively engaged in cognitively demanding activities and the extent to which they developed a community of learners in which all students could learn from each other as well as from the teacher. It measured the extent to which teachers provided their students with opportunities to develop an understanding of how mathematical concepts are related and encouraged diverse thinking and multiple solution strategies. The last scale, the Multiple Representations scale, captured the extent to which teachers and students used a variety of manipulatives or representations to develop the concepts.

Data analysis

To investigate the extent to which teachers' instructional practices were improved over the duration of the program, we used multilevel growth modeling. Growth modeling is one of the most appropriate approaches for analyzing longitudinal data sets because it is specifically designed to study change. In this study, growth modeling was used to investigate how teachers' practices improved throughout their participation in professional development activities. Given that the participating teachers were observed up to seven times in 4 years, this method allowed us to characterize their improvement in instruction over time. In addition, because changes in instruction could be either linear (i.e., increasing steadily over time) or nonlinear (i.e., increasing for a while and then decreasing), growth modeling allowed us to capture trends in the data more accurately. Furthermore, the differences in teachers' initial practices could be adjusted based on differences in their backgrounds. Moreover, unlike repeated measures, growth modeling could handle missing data efficiently (Hedeker 2004; Hedeker and Gibbons 2006).

Given that we had data on teachers' mathematics instruction from different time points and that teachers were nested within cohorts, we used a three-level multivariate model (see Fig. 1). More specifically, in Level 1 (time), for each of the five scales that emerged from the factor analysis, individual teachers' trajectories of change on that scale were predicted based on their years in the program and teaching experience.[7] In Level 2 (teachers), differences in teachers' baseline practices were predicted based on their mathematics content knowledge, as measured by a test designed by the center. In Level 3, the teachers' cohort was adjusted because teachers from Cohort 1 and Cohort 2 attended different summer institutes.[8] To address the second research question regarding which instructional practices were more apt to change and which were more resistant, we standardized each instructional practice scale score by using teachers' initial practices so that change could be interpreted via standardized regression for the five scales and the comparison would be more meaningful.[9]

Results

Table 2 presents descriptive statistics for the teachers' initial scores on the scales when their first observations were conducted. As shown, based on practices captured by our classroom observation instruments, the teachers had ample room to improve their practices.

[7] Teachers' grade levels were not included in the analysis because many teachers were teaching multiple grade levels in a given year.

[8] Analyses were also conducted separately for the two cohorts to investigate whether the overall observed trend was similar for both cohorts, which was similar across all five scales. In addition, an alternative two-level model with cohort added as a Level 2 variable resulted in similar coefficients for the "time" variable.

[9] All models were fit using the mixed methods procedure (PROC MIXED) in SAS/STAT software (SAS Institute Inc., 2008). Additionally, given that the data were collected from the same teachers over time, residuals within teachers were correlated. Hence, the autocorrelated error needed to be taken into consideration in addition to the random error (e.g., measurement errors and missing variables; Hedeker and Gibbons 2006). The autoregressive structure was used to test whether a serial correlation needed to be taken into account.

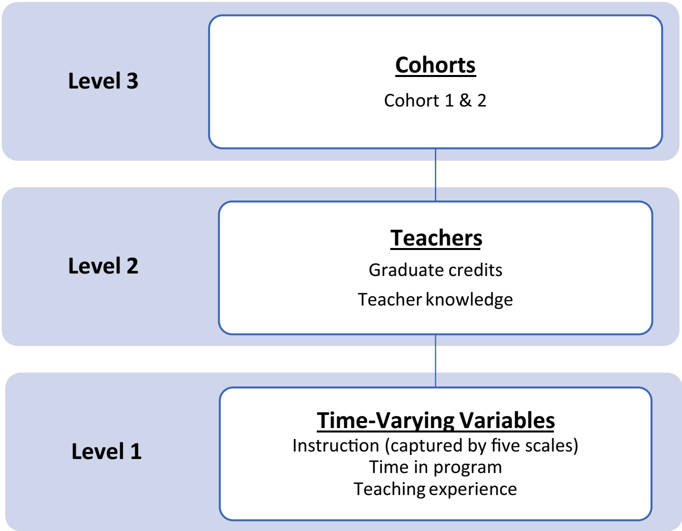


Fig. 1 Levels in the multilevel modeling

Table 2 Initial status of teachers' scores on the outcome scales

Scales	Mean	SD	Minimum	Maximum
Student Interactions	2.37	1.43	0	5
Mathematical Discourse	2.99	1.24	0	5
Instructional Clarity	2.70	1.66	0	5
Mathematical Habit of Mind	0.59	0.32	0	1
Multiple Representations	0.46	0.41	0	1

Changes in teachers' instructional practices

As shown in Table 3, teachers' scores on four out of five scales[10] increased significantly throughout the program.[11] On average, the growth on the Mathematical Discourse, Instructional Clarity, and Mathematical Habit of Mind scales was linear. Specifically, teachers' scores on the Mathematical Discourse and Mathematical Habit of Mind scales increased by .15 and .11 SD per semester ($p \leq .0001$ and $p = .003$, respectively). This result indicates that teachers continued to improve the quality of their mathematical discourse and provided more opportunities for their students to develop mathematical habits of mind in which the students were investigating each other's ideas, working on mathematically challenging problems, and seeing the connections between what was learned and other mathematical ideas and real-world applications. Unlike the linear growth pattern observed in the aforementioned scales,

[10] The residual correlation was not significant for any of the models for all scales, which indicates that the correlation between errors was not significant. Neither teachers' scores on the content knowledge instrument nor years of teaching experience were related to the improvement in their mathematics instruction; therefore, they were excluded from the data analysis and from the final models reported here.

[11] Teachers' scores on the Multiple Representations scale did not change noticeably throughout the program ($p = .30$).

Table 3 Standardized regression coefficients and standards errors for growth models for the instructional practice scales

Variables	Student Interactions	Mathematical Discourse	Instructional Clarity	Mathematical Habit of Mind	Multiple Representations
Teachers' average practices in the first fall semester in the program	−.06 (.24)	.01 (.26)	.05 (.58)	−.01 (.26)	−.09 (.15)
Average rate of change in teachers' practices by semester	0.42*** (.11)	0.15*** (.04)	0.07~ (.04)	0.11** (.04)	0.03 (.03)
Semester^2 (an indicator that the change in teachers' practices followed a quadratic pattern)	−0.06** (.02)				

Standardized regression coefficients are shown, with standard errors in parentheses

*** $p < .0001$; ** $p < .001$; * $p < .05$; ~ $p \leq .10$

the change in Student Interactions followed a quadratic pattern.[12] Teachers first began to make drastic changes, as captured by this scale (see the coefficient for semester and semester^2 in Table 3), but they were not able to sustain these changes at the same level.[13] This result indicates that the teachers first provided time and opportunities for their students to learn from each other and that the students did work together and did explain their ideas to their peers; however, the teachers then began to return to their original practices by providing less time and fewer opportunities for their students to learn from each other. Teachers' predicted scores on the Instructional Clarity scale increased by .07 SD with each subsequent measurement ($p = .052$).

Sustainable changes in instructional practices

In the previous section, we reported the extent to which teachers' scores on the five instructional practice scales changed over time. Here, we compare which of these practices were more apt to change. To do so, we created Figs. 2 and 3, which represent the predicted changes over time based on the data. As shown in Fig. 2, teachers made drastic changes in student interactions, although these changes were not retained by the end of the program. On the other hand, teachers were able to continue improving their practices in the areas of mathematical discourse and instructional clarity. Additionally, changes in their quality of mathematical discourse were more substantial than those made in their instructional clarity. The total change in the Mathematical Discourse scale at the end of the program was more than 1.06 SD, whereas the total change in the Instructional Clarity scale was .54 SD. As can be seen in Fig. 3, teachers' scores on the Mathematical Habit of Mind scale also increased gradually, and the total change was .76 SD at the end of the program. On the other hand, teachers' scores on the Multiple Representations scale did not change noticeably.

[12] The coefficient for semester^2 in Table 3 represents a quadratic change on the Student Interactions scale.

[13] Because the coefficient for semester was positive ($p < .001$) but for semester^2 was negative ($p < .002$), teachers' scores on the Student Interactions scale initially increased and then gradually decreased.

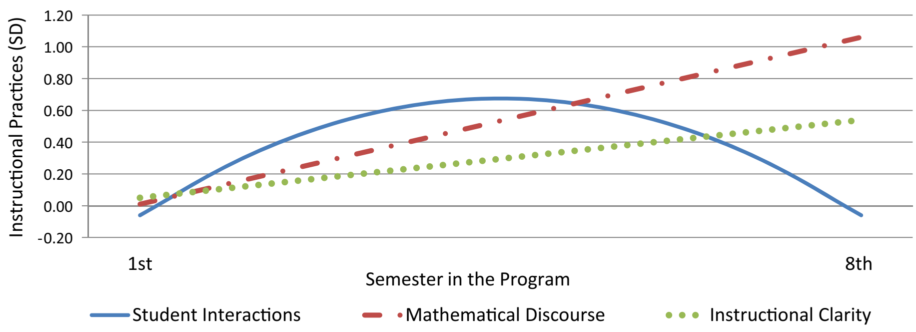


Fig. 2 Growth curves for the Student Interactions, Mathematical Discourse, and Instructional Clarity scales

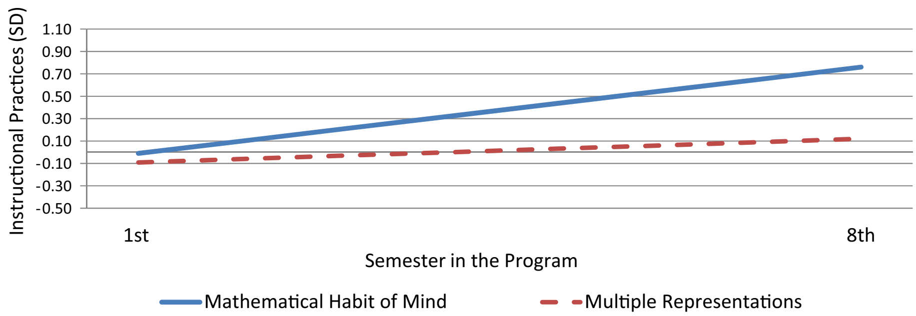


Fig. 3 Growth curves for the Mathematical Habit of Mind and Multiple Representations scales

Comparisons of teachers' scores showed discernible patterns in their instructional growth over the years. Teachers seemed not to sustain changes in their practices toward providing students with opportunities to learn from each other over time, and they did not make any modifications in their instruction to incorporate multiple representations. Furthermore, of the three practices that changed gradually, the most noticeable change was their improvement in the mathematical quality of classroom discourse, then in creating a community of learners in which all students and the teacher worked together to build mathematical understandings by solving complex mathematical problems (Mathematical Habit of Mind), and finally by making explicit to the students the expectations and goals of the lessons and activities. It seems, based on these results, that teachers had difficulty clarifying the instructional objectives but could relatively more easily create an environment in which students could share, discuss, and justify their solutions.

Discussion

Results of this study indicated that the teachers changed several aspects of their instruction after they had attended the professional development program. They made prominent and permanent changes in their practices so that students shared their knowledge, justified their conclusions, and talked more about mathematics (Mathematical Discourse). Similarly, the students actively engaged in thought-provoking activities, and the teachers created an environment that allowed the students to become members of a learning community in which intellectual rigor and constructive criticism were valued (Mathematical Habit of Mind). The teachers also became more explicit regarding what they expected from their students during group activities and whole-class discussions (Instructional Clarity). However, the teachers did not show steady growth in practices that would allow their students time to discuss and share what they understood with their peers, to work in groups, or to explain their understandings to their peers (Student Interactions). Finally, neither the teachers nor students began to use a variety of means to present concepts after the program was initiated (Multiple Representations). In the following sections, we interpret the findings from three different angles.

Key features of the program

The findings of this study support earlier work showing that content-focused professional development programs, by actively engaging teachers in meaningful mathematical activities, had a sustainable effect on teachers' classroom practices (e.g., Boston and Smith 2009; Farmer et al. 2003; Franke et al. 2001; Parise and Spillane 2010; Supovitz et al. 2000). The activities in the program were explicitly designed to promote discourse by immersing the participants in problem-solving experiences to deepen their understanding of mathematics content. In addition, we believe that two key features of effective professional development contributed to this impact on teachers' practices: (1) its connections to teachers' work and (2) the duration of the program.

As several scholars have noted, the success of teacher change is dependent on the extent to which the content of professional development activities is aligned with the teachers' daily work of teaching and with state and district standards (e.g., Borko et al. 1997; Desimone 2009; Klein 2001). The staff at the center who were part of the design and implementation of the professional development program were also teachers at one point in their careers, so they had a strong understanding of the importance of connecting all activities to teachers' work and to state and district standards. In addition, adjusting the topics of the academic-year meetings based on the needs of the participating teachers seemed to help teachers receive immediate help and overcome the pressure and dilemmas they were facing while implementing reform-based instruction.

The second vital component was the duration of the program. The findings of this study echo those of earlier studies showing that professional development programs with a greater number of contact hours and of long duration have a positive impact on teachers (e.g., Garet et al. 2001; Supovitz and Turner 2000). This result also contributes to the existing literature by providing evidence of the types of follow-up activities that promote such profound and sustainable changes in teachers' practices. We believe that assigning each teacher a mentor and visiting participating teachers' classrooms to observe their practices enabled the professional development organizers to identify the specific needs of each individual teacher and address them more efficiently. In addition, recall that some teachers attended the academic-year meetings for more than 4 years and had opportunities to learn from other teachers as well as the center personnel. This ongoing support over the years made it possible for teachers to show continued improvement in their mathematics instruction.

Practices that could be improved by professional development programs

Results of the comparison of teachers' change patterns over the years suggest that certain practices were apt to change, whereas others were not. For instance, changes in teachers' practices captured on the Student Interactions scale seemed to be easy to change but difficult to sustain. According to these findings, teachers changed their practices so that their students would work in groups and explain their answers, and they did preserve these changes. In the summer institutions, the teachers were encouraged to let their students work in groups and provide explanations for their responses and conjectures, so the teachers first seemed to be motivated to make changes in these practices. However, as time passed, the teachers moved away from these practices, and by the end of the program, their practices were not so different from their original ones. Our earlier work with different groups of teachers at the elementary and middle school levels also indicated that instructional practices related to student participation were harder to sustain because each year, teachers had a different body of students, which affected their instructional choices, especially those choices pertaining to allowing students to interact with one another (Copur-Gencturk 2012; Copur-Gencturk et al. 2014).

One of the most encouraging findings of this study was that teachers continued to improve the quality of their mathematical discourse, the clarity of their instruction and activities, and the mathematical quality and cognitive demands of their mathematics instruction. The drastic and continued improvements in these scales suggest that ongoing and individualized support could be an efficient follow-up activity to help sustain changes the teachers had made.

Finally, the findings of this study suggest that teachers were hesitant to incorporate multiple representations into their teaching in spite of the program's encouragement and emphasis on doing so. This result is not surprising considering that US high school teachers reported limited incorporation of hands-on activities, especially in the advanced mathematics courses (Smith 2013) even though more than half of these teachers reported attending professional development activities that placed heavy emphasis on how to incorporate hands-on activities or manipulatives into their mathematics instruction. Although research supports the claim that providing opportunities for students to work with visual representations or concrete materials helps students understand mathematics and become more effective problem solvers (Gersten et al. 2014; Woodward et al. 2012), a lack of time and available resources could be the reasons why teachers preferred not to use multiple representations. Further studies should aim to identify why high school teachers were reluctant to incorporate multiple representations into their mathematics instruction.

Limitations

The study has several limitations. First, the teachers in this study were not a random sample of high school teachers. Further studies that investigate the generalizability of the observed patterns are needed. In addition, the sample size for teachers may have limited the possibility of finding some existing relationships. However, in a multilevel longitudinal analysis, sample size should be considered separately for each level; thus, increasing the number of teachers might not increase the power of analysis as much as would increasing the number of time points. Because we were interested in the change in teachers' practices, having more time points increased the power of the study more than did having a greater number of teachers.

A second potential limitation is related to the instrument used in the study. Although the external evaluator of the project has expertise in instrument development and the center personnel discussed the items and reached agreement on what each item was designed to measure, no established validity and reliability scores were available. All measures, even well-established ones, included some measurement error, which reduced the possibility of finding and measuring relationships accurately. Thus, the observed relationships might be affected by measurement error. Finally, another potential source of measurement error was related to the limited number of observations conducted in teachers' classrooms. A greater number of classroom observations might reflect teachers' "typical" mathematics instruction, hence capturing teachers' practices more accurately.

Implications

This study suggests that several key features of high-quality professional development can indeed help teachers improve their teaching practices. Focusing on the content the participating teachers need to enhance their understanding of mathematics instruction, creating opportunities for the teachers to actively build on their knowledge, connecting what is taught in the teachers' professional development activities with their actual work, and providing continuing support not only can improve the teachers' instruction, but also can help them sustain those changes. Of the four key features of effective professional development, the importance of appropriate, ongoing support is the feature highlighted in this study. Teachers continued to improve their practices after they had completed the summer institutes and during the follow-up activities. Results of this study inform teacher educators that ongoing support not only helps teachers maintain the changes they have made in their practices, but also improves their practices further. Therefore, the findings of this study highlight the importance of specific follow-up activities utilized in the program, such as observing teachers and providing feedback on their teaching, as well as meeting regularly to address issues raised in their teaching.

The study informs researchers and teacher educators about which aspects of instructional practices are more apt to change and which are more resistant to change. Earlier studies looked at teachers' instruction holistically and assumed that changes in various aspects of their practices followed the same trends. The findings of the present study challenge this assumption and urge researchers to conduct more studies to identify which features of mathematics instruction could be improved through professional development activities.

Acknowledgments This study is based upon work sponsored by a grant from the National Science Foundation under the Grant No. EHR 0412072. Any opinions, findings, and conclusions expressed herein are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Abell, S. K., & Pizzini, E. L. (1992). The effect of a problem solving in-service program on the classroom behaviors and attitudes of middle school science teachers. *Journal of Research in Science Teaching, 29*(7), 649–667.

American Federation of Teachers. (2002). *Principles for professional development: AFT's guidelines for creating professional development programs that make a difference*. Washington, DC: Author.

Anthony, G., & Walshaw, M. (2009). Characteristics of effective teaching of mathematics: A view from the West. *Journal of Mathematics Education, 2*(2), 147–164.

Avalos, B. (2011). Teacher professional development in teaching and teacher education over ten years. *Teaching and Teacher Education, 27*(1), 10–20.

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

Banilower, E. R., Heck, D. J., & Weiss, I. R. (2007). Can professional development make the vision of the standards a reality? The impact of the National Science Foundation's local systemic change through teacher enhancement initiative. *Journal of Research in Science Teaching, 44*(3), 375–395.

Bell-Gredler, M. E. (1986). Jean Piaget's cognitive-development theory. *Learning and instruction: Theory into practice* (pp. 191–233). New York: Macmillan.

Blank, R. K., & de las Alas, N. (2010). *Effects of teacher professional development on gains in student achievement: How meta analysis provides scientific evidence useful to education leaders*. Washington, DC: Council of Chief State School Officers

Borko, H., Mayfield, V., Marion, S., Flexer, R., & Cumbo, K. (1997). Teachers' developing ideas and practices about mathematics performance assessment: Successes, stumbling blocks, and implications for professional development. *Teaching and Teacher Education, 13*(3), 259–278.

Boston, M. D., & Smith, M. S. (2009). Transforming secondary mathematics teaching: Increasing the cognitive demands of instructional tasks used in teachers' classrooms. *Journal for Research in Mathematics Education, 40*(2), 119–156.

Boston, M. D., & Smith, M. S. (2011). A 'task-centric approach' to professional development: Enhancing and sustaining mathematics teachers' ability to implement cognitively challenging mathematical tasks. *ZDM, 43*(6–7), 965–977.

Bullough, R. V., Kauchak, D., Crow, N. A., Hobbs, S., & Stokes, D. (1997). Professional development schools: Cataylsts for teacher and school change. *Teaching and Teacher Education, 13*(2), 153–169.

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

Copur-Gencturk, Y. C. (2012). *Teachers' mathematical knowledge for teaching, instructional practices, and student outcomes* (Doctoral dissertation), University of Illinois at Urbana-Champaign.

Copur-Gencturk, Y. (2015). The effects of changes in mathematical knowledge on teaching: A longitudinal study of teachers' mathematical knowledge and instruction. *Journal for Research in Mathematics Education, 46*(3), 280–330.

Copur-Gencturk, Y., Hug, B., & Lubienski, S. T. (2014). The effects of a master's program on teachers' science instruction: comparing classroom observations, teacher reports, and student surveys. *Journal for Research in Science Teaching, 51*(2), 219–249.

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher, 38*(3), 181–199.

Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis, 24*(2), 81–112.

Desimone, L., Smith, T. M., & 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.

Dewey, J. (1997). *Experience and education*. New York: Touchstone **(Original work published 1938)**.

Farmer, J. D., Gerretson, H., & Lassak, M. (2003). What teachers take from professional development: Cases and implications. *Journal of Mathematics Teacher Education, 6*(4), 331–360.

Fishman, B. J., Marx, R. W., Best, S., & Tal, R. (2003). Linking teacher and student learning to improve professional development in systemic reform. *Teaching and Teacher Education, 19*(6), 643–658.

Franke, M. L., Carpenter, T. P., Levi, L., & Fennema, E. (2001). Capturing teachers' generative change: A follow-up study of professional development in mathematics. *American Educational Research Journal, 38*(3), 653–689.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*(4), 915–945.

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). Washington, DC:

U.S. Department of Education, Institute of Education Sciences, National Center for Educational Evaluation and Regional Assistance, Regional Educational Laboratory Southeast.

Gibbons, S., Kimmel, H., & O'Shea, M. (1997). Changing teacher behavior through staff development: Implementing the teaching and content standards in science. *School Science and Mathematics, 97*(6), 302–310.

Guskey, T. R. (2002). Does it make a difference? Evaluating professional development. *Educational Leadership, 59*(6), 45–51.

Guskey, T. R. (2003). Analyzing lists of the characteristics of effective professional development to promote visionary leadership. *NASSP Bulletin, 87*(637), 4–20.

Guskey, T. R., & Yoon, K. S. (2009). What works in professional development. *Phi Delta Kappan, 90*(7), 495–500.

Hedeker, D. (2004). An introduction to growth modeling. In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social sciences* (pp. 215–234). Thousand Oaks, CA: Sage.

Hedeker, D. R., & Gibbons, R. D. (2006). *Longitudinal data analysis*. New York: Wiley.

Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction, 26*(4), 430–511.

Ingvarson, L., Meiers, M., & Beavis, A. (2005). Factors affecting the impact of professional development programs on teachers' knowledge, practice, student outcomes & efficacy. *Educational Policy Analysis Archives, 13*(10), 1–26. Retrieved January 2, 2013, from http://epaa.asu.edu/epaa/v13n10/.

Klein, B. S. (2001). Guidelines for effective elementary science teacher in-service education. *Journal of Elementary Science Education, 13*(2), 29–40.

Lambert, L. (1998). *Building leadership capacity in schools*. Alexandria, VA: ASCD.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal, 27*(1), 29–63.

Lee, O., Hart, J. E., Cuevas, P., & Enders, C. (2004). Professional development in inquiry based science for elementary teachers of diverse student groups. *Journal of Research in Science Teaching, 41*(10), 1021–1043.

Leinwand, S. (2000). *Sensible mathematics: A guide for school leaders*. Portsmouth, NH: Heinemann.

Loucks-Horsley, S., & Matsumoto, C. (1999). Research on professional development for teachers of mathematics and science: The state of the scene. *School Science and Mathematics, 99*(5), 258–271.

Loucks-Horsley, S., Stiles, K. E., Mundry, S., Hewson, P. W., & Love, N. (Eds.). (2010). *Designing professional development for teachers of science and mathematics*. Thousand Oaks, CA: Corwin Press.

McMeeking, L. B. S., Orsi, R., & Cobb, R. 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.

Merrill, C., Devine, K. L., Brown, J. W., & Brown, R. A. (2010). Improving geometric and trigonometric knowledge and skill for high school mathematics teachers: A professional development partnership. *Journal of Technology Studies, 36*(2), 20–30.

Miller, B., Moon, J., & Elko, S. (2000). *Teacher leadership in mathematics and science: Casebook and facilitator's guide*. Portsmouth, NH: Heinemann.

Moyer-Packenham, P. S., Bolyard, J. J., Oh, H., & Cerar, N. I. (2011). Common features of professional development activities for mathematics and science teachers. *Professional Development in Education, 37*(4), 571–589.

National Academy of Education. (2009). *Teacher quality*. Washington, DC: Author.

National Council of Teachers of Mathematics [NCTM]. (1991). *Professional standards for teaching mathematics*. Reston, VA: NCTM.

NCTM. (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.

NCTM. (2014). *Principles to actions: Ensuring mathematical success for all*. Reston, VA: NCTM.

Parise, L. M., & Spillane, J. P. (2010). Teacher learning and instructional change: How formal and on-the-job learning opportunities predict change in elementary school teachers' practice. *The Elementary School Journal, 110*(3), 323–346.

Penuel, W. R., Fishman, B. J., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal, 44*(4), 921–958.

Polly, D., & Hannafin, M. J. (2011). Examining how learner-centered professional development influences teachers' espoused and enacted practices. *The Journal of Educational Research, 104*(2), 120–130.

Sandholtz, J. H. (2002). Inservice training or professional development: Contrasting opportunities in a school/university partnership. *Teaching and Teacher Education, 18*(7), 815–830.

Sandholtz, J. H., & Ringstaff, C. (2013). Assessing the impact of teacher professional development on science instruction in the early elementary grades in rural US schools. *Professional Development in Education, 39*(5), 678–697.

SAS Institute, Inc. (2008). *SAS/STAT1 9.2 user's guide.* Cary, NC: SAS Institute.

Scher, L., & O'Reilly, F. (2009). Professional development for K-12 math and science teachers: What do we really know? *Journal of Research on Educational Effectiveness, 2*(3), 209–249.

Schoen, H. (Ed.). (2003). *Teaching mathematics through problem solving, grades 6–12*. Reston, VA: NCTM.

Senk, S. L., & Thompson, D. R. (Eds.). (2003). *Standards-based school mathematics curricula: What are they? What do students learn?*. Mahwah, NJ: Lawrence Erlbaum.

Skemp, R. R. (1987). *The psychology of learning mathematics*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Smith, A. A. (2013). *Report of the 2012 national survey of science and mathematics education: Status of middle school mathematics*. Chapel Hill, NC: Horizon Research.

Supovitz, J. A., Mayer, D. P., & Kahle, J. B. (2000). Promoting inquiry-based instructional practice: The longitudinal impact of professional development in the context of systemic reform. *Educational Policy, 14*(3), 331–356.

Supovitz, J. A., & Turner, H. M. (2000). The effects of professional development on science teaching practices and classroom culture. *Journal of Research in Science Teaching, 37*(9), 963–980.

Van Hiele, P. M. (1986). *Structure and insight: A theory of mathematics education*. Orlando, FL: Academic Press.

Van Hiele, P. M., & Van Hiele-Geldof, D. (1958). A method of initiation into geometry at secondary school. In H. Freudenthal (Ed.), *Report on methods of initiation into geometry* (pp. 67–80). Groningen: J. B. Wolters.

Vygotsky, L. S. (1962). In E. Hanfmann & G. Vakar (Eds.), *Thought and language* (E. Hanfmann & G. Vakar, Trans.). Cambridge, MA: Massachusetts Institute of Technology.

Weiss, I. R., Pasley, J. D., Smith, P. S., Banilower, E. R., & Heck, D. J. (2003). *Highlights report, looking inside the classroom: A study of K-12 mathematics and science education in the United States*. Chapel Hill, NC: Horizon Research.

Whittington, D. (2002). *Report of the 2000 national survey of science and mathematics education: Status of high school mathematics teaching*. Chapel Hill, NC: Horizon Research.

Woodward, J., Beckmann, S., Driscoll, M., Franke, M., Herzig, P., Jitendra, A., et al. (2012). *Improving mathematical problem solving in grades 4 through 8* education. Retrieved from http://ies.ed.gov/ncee/wwc/publications_reviews.aspx#pubsearch/.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*(4), 458–477.