Context-driven learning from professional development: Examining teacher knowledge, instructional practice, and student mathematics achievement gains

Jennifer K. Jacobs
University of Colorado Boulder

Karen Koellner
Hunter College, The City University of New York

Nanette Seago
WestEd

Helen Garnier
Consultant

Chao Wang
University of Colorado Boulder

AUTHOR NOTES:
The Learning and Teaching Geometry Efficacy Study was supported by the National Science Foundation (NSF award #1503399).

To be presented at the American Educational Research Association, April 2019, Toronto Canada.

DRAFT PLEASE DO NOT DISTRIBUTE OR CITE WITHOUT PERMISSION OF THE AUTHORS

Abstract

Professional development (PD) that sets ambitious goals for teacher learning leading to instructional change is in high demand. Even more critical, but somewhat elusive, are PD programs that have a demonstrable effect on teacher knowledge and classroom practice. Supporting teachers to skillfully implement the Common Core State Standards is currently a key focus of many mathematics PD efforts, with an eye towards helping teachers gain mathematical knowledge for teaching and fluency in instructional decision-making that support students' learning of challenging content. The Learning and Teaching Geometry (LTG) intervention consists of well-specified PD materials that engage teachers in learning complex geometric concepts targeted in the standards through viewing and discussing videocases. This paper describes a group-randomized study to examine the effectiveness of the LTG intervention in two geographic locations. Although the impact of the intervention on teachers' knowledge was modest, a stronger effect on their instructional practice emerged which then predicted greater student gains in the targeted content area. Additionally, teachers in the two locations exhibited context-based differences in their learning. This paper explores the gains teachers made, along with possible explanations for the differences between the locations and implications for future research.

Context-driven learning from professional development: Examining teacher knowledge, instructional practice, and student mathematics achievement gains

Theoretical Framework and Objectives

The mathematics education literature indicates that there are consistent and positive links between teachers' mathematical knowledge, instructional quality, and student outcomes (e.g., Desimone, 2009; Jacob, Hill & Corey, 2017). A commonly accepted theoretical trajectory of teacher learning suggests the following pathway: gains in teacher knowledge lead to subsequent changes in classroom instruction which, in turn, support improved student learning (Desimone, 2009; Yoon et al., 2007). In other words, the impact of increased teacher knowledge on student knowledge growth is thought to be mediated by improved instructional quality (Kersting, Givvin, Sotelo & Stigler, 2010). There is some preliminary evidence that this hypothesized trajectory holds true, based on studies that have measured all three elements (Baumert et al., 2010; Kersting et al., 2012).

A central challenge for the field of teacher professional development is how to design interventions that target teacher knowledge, while also maintaining a focus on instructional practice and student learning. A number of researchers have worked to address this challenge and there is now a strong research base delineating critical design aspects of effective PD (e.g., Borko, Jacobs & Koellner, 2010; Desimone et al., 2002; Ingvarson, Meiers & Beavis, 2005; Penuel, Fishman, Yamaguchi & Gallagher, 2007). At the same time, studies of PD outcomes yield a mixed picture. Although some PD programs that adhere to design recommendations by the literature have produced encouraging results (e.g. Franke, Carpenter, Levi & Fennema, 2001; Kutaka et al., 2017; Taylor et al, 2017), others have proven much less successful (e.g. Jacob, Hill & Corey, 2017; Santagata, Kersting, Givvin & Stigler, 2010). Much remains to be learned about the development and implementation of effective PD.

The Value of Video-based PD

Designing professional development that incorporates classroom video has become increasingly popular, as numerous studies have documented the potential of viewing and discussing video to foster teacher learning. Two recent, comprehensive literature reviews highlight the substantial promise of video-based PD (Gaudin & Chalies, 2015; Major & Watson, 2018). Gaudin and Chalies (2015) reviewed 255 articles related to video viewing in teacher education and PD, concluding that classroom video is an especially productive learning tool with regards to improving instructional practice. They argue, "…the value of video use by teachers lies principally in the opportunity to raise teachers' quality of instruction" (pg. 59). Major & Watson's (2018) limited their review to the use of video in PD programs serving in-service teachers. Based on 82 studies, they concluded that not only is video a powerful and effective tool to promote improvements in classroom practice, it has the potential to enhance teacher cognition as well.

When purposefully selected and integrated into a carefully designed PD framework, video footage from classroom lessons can advance specified learning goals, such as the improvement of particular instructional practices (Blomberg et al. 2013). Targeted and scaffolded video viewing engages teachers in reflection, providing the opportunity to "mentally simulate instructional action" (Blomberg et al., 2014, pg. 445). Engaging in this type of self-reflective process while viewing video contributes to the development of teachers' "professional vision," a type of knowledge that includes the ability to notice and make sense of events in an instructional setting (Sherin, 2007; Sherin, Jacob & Philipp, 2011; Sherin & Van Es, 2009). Van Es and Sherin (2010) found that as teachers become more adept at analyzing videos through the lens of student thinking, they carried this focus into their classrooms in order to elicit, probe, and build on students' ideas. Similarly, research by Kersting and colleagues (Kersting, 2008; Kersting et al., 2012) suggests that teachers' ability to analyze and interpret video clips of mathematics classrooms is positively related to the instructional quality of their own mathematics lessons.

Video Promotes Community Formation and Collaborative Inquiry

Within a professional development context, videos of classrooms can be understood as boundary objects (Kazemi & Hubbard, 2008) in the sense that the videos help to maintain coherence across multiple and intersecting social worlds (Star & Griesmer, 1989). In their function as a boundary objects, videos enable a shared space within which the teachers and PD facilitator can form a professional community. As such, videos serve as analytical sites that enable multiple viewpoints and ideologies to surface, from multiple parties (Miller & Zhou, 2007). Furthermore, teachers have a dual role in PD settings; they are both "teachers" and "students" or learners. The analysis of video can support teachers to take up these roles in specified moments, and traverse boundaries as needed. Video can encourage "perspective taking" (Akkerman & Bakker, 2011) by helping teachers see themselves in both the teacher and student role, and reflexively understand the needs and limitations of these two identities.

Although viewing video is in many ways a self-directed exercise in which teachers attend to topics of their own interest and construct personally relevant knowledge, when used as a boundary object in a collaborative PD setting, video viewing can encourage increasingly productive conversations and community building (Borko et al., 2008; van Es & Sherin, 2008). Using video as a tool to support collaborative inquiry and engagement can anchor dialogue and lead to collective ideas for improvement (Koh, 2015). For example, van Es (2012) reports that in the video club PD meetings she analyzed, over time the teachers began to hold more substantive discussions about specific problems of practice that were pertinent to the group members, and the community adopted more collegial discourse norms.

Situative Theory and the Intersection of Video, Community, and Learning

Situative theory suggests that learning is situated in a particular context, socially organized, and distributed across individuals, artifacts and tools (Putnam & Borko, 2000). Learning emerges through authentic activities and social interactions within a community (Brown, Collins, & Duguid, 1988; Lave & Wenger, 1991). The learning setting plays a critical role, and each learning setting is unique with regards to the types of knowledge and experiences that the participants bring with them (Bransford, Brown & Cocking, 2000).

Using classroom video in PD situates teacher learning firmly within their everyday routines of practice. In addition, the task of analyzing and making sense of video presents a high degree of complexity, providing teachers with motivation to socially engage and interact cooperatively (Blomberg et al., 2013). Hatch, Shuttleworth, Jaffee and Marri (2016) propose three elements that shape teachers' learning from video in a professional development setting: (1) properties of the videos, (2) viewers' knowledge and experiences, and (3) social factors and activities in which videos are examined. They argue, "…the situated perspective highlights that the affordances for learning from video emerge from the interaction of all three" (pg. 276). In other words, situative theory helps make clear that teacher learning occurs as a result of the relationship between selected video representations of teaching, participants' unique backgrounds, and socially organized viewing experiences.

How Might Context Influence What Teachers Do and Learn Within PD?

A situative perspective suggests that groups of teachers who take part in PD workshops using the same materials, with the same facilitator, but situated within different educational contexts (e.g., different geographical locations within the United States) might have very different learning opportunities and experiences. Even within a highly specified video-based PD, in which resources and facilitation materials are provided to ensure a particular experience (Koellner & Jacobs, 2015), the nature of teachers' conversations and the degree to which their learning is impacted is unlikely to be constant across contexts. For example, precisely what teachers notice about the video clips can yield unique perspectives, conversations, and reflections within a given group of participants. The knowledge and prior experiences teachers bring with them to the workshops is also likely to be quite different across locations. Additionally, as noted previously, during PD workshops the participating teachers develop a unique professional community, which serves as the social unit within which learning takes place. As Webster-Wright (2009) argued, "Learning always occurs in a context, as has been highlighted by the range of research into the social, situated nature of learning." (p. 723).

At present very little is known about the degree to which context impacts teachers' learning from professional development. Do teachers from different locations who participate in the same PD program with the same facilitator experience differential impacts on their knowledge, practice, and student learning? We argue there are at least three reasons to conjecture that the answer to this question is likely to be affirmative: (1) wide variation exists across geographic locations (i.e. US states) on nearly all educational measures, such as teacher-pupil ratios, teacher salaries, teacher experience levels, teacher turnover, student achievement, and professional development opportunities (Grissmer et al., 2000; Wei, Darling-Hammond & Adamson, 2010); (2) schools and districts present distinct organizational and workplace climates that affect teachers' formal and informal learning and growth opportunities (Avalos, 2011; Cobb, McClain, Lamberg, & Dean, 2003; Firestone et al., 2005); and (3) teachers form unique professional learning communities within PD contexts that impact their collective learning and individual growth (Grossman, Wineburg & Woolworth, 2001; Little, 2003).

Borko (2004) argues for the importance of research on well-specified PD programs implemented in multiple sites with multiple facilitators, in order for the field to move beyond "existence proofs" of effective PD. In our study, we implemented a video-based mathematics PD program in two unique geographic locations within the US. Both sites were facilitated by the same individual using highly detailed and extensive facilitation materials to ensure consistency in implementation. As such, our research joins a very limited body of literature that is able to compare the effectiveness of a given PD program across multiple contexts. Yet, even widely scaled professional learning endeavors such as the National Writing Project (Academy for Educational Development, 2002; Gallagher, Woodworth & Arshan, 2015) tend not to disaggregate their data across sites to examine location effects, particularly in a quantitative fashion. Even when research on PD is conducted and reported across sites, variation in effectiveness is generally attributed to differences in facilitation or other aspects of the implementation of the PD (e.g., Bell, Wilson, Higgins & McCoach, 2010).

The Need for Professional Learning in the Area of Geometric Transformations

A critical feature of effective PD is that it addresses a problem of practice, meaning that it meets the professional needs of teachers (Scribner, 1999). Starting in middle school, the Common Core State Standards for Mathematics (CCSSM) contain a strong and consistent focus on geometric transformations—including their mathematical properties, how they can be sequenced, and their effect on two-dimensional figures in a coordinate plane. For example, seventh graders are expected to understand how to scale drawings of geometric figures. The eighth-grade standards encourage an in-depth understanding and application of geometric transformations, along with connections between proportional reasoning, slope, and linearity to geometric transformations and similarity. The high school standards continue to promote a transformations-based approach to understanding congruence, similarity, and the proof of various relationships among geometric figures based on similarity. This increased emphasis on transformational geometry represents a major difference from previous state standards (Teuscher, Reys & Tran, 2016; Tran et al., 2016).

As Teuscher and colleagues (2016) explain, "Most high school students are taught that for two figures to be similar they must have corresponding congruent angles and corresponding side lengths proportional. This again brings up the question, why? With an understanding of geometric transformations, students can answer the why question as they make connections between these concepts." (pg. 11). Several reports have called on professional developers to create materials that emphasize transformations-based geometry, in accordance with the CCSSM (McCallum, 2011; Sztajn, Marrongelle, Smith & Melton, 2011). A national task force of mathematicians, mathematic educators, state leaders, and teacher leaders identified Grade 8 Geometry, with an emphasis on transformations, as one of five priority areas for CCSS mathematics PD and resources in grades K-8 (McCallum, 2011). PD materials that help support teachers in these priority areas can play a timely role in the present mathematics education environment.

Teachers generally agree that they need much more support to learn and effectively implement the Common Core State Standards in their classrooms. In an online survey conducted by the EPE Research Center, a majority of teachers responded that they had participated in some PD related to the CCSS and agreed that the standards would help them improve their classroom practice (Gewertz, 2013). However, across several studies teachers reported only a moderate level of preparedness to teach the standards to their students, and they indicated a strong desire for professional learning opportunities to ensure alignment between their curricular materials and the standards (Gewertz, 2013; Perry et al., 2015; Roth McDuffie et al., 2017).

The Learning and Teaching Geometry PD Intervention

The Learning and Teaching Geometry (LTG) program (NSF study award #0732757) is a video-based mathematics professional development intervention, targeted for teachers serving grades 6-12. The intervention consists of 54 hours of PD focused on improving teaching and learning of mathematical similarity based on geometric transformations (Seago et al., 2017). The program is designed to be implemented by a knowledgeable facilitator, using a set of provided resources to engage teachers in a specified learning trajectory aligned with multiple middle and high school Common Core State Standards (Seago et al., 2013). The intervention includes a 30-hour Foundation Module followed by four 8-hour Extension Modules that explore related topics such as using appropriate representations and tools and supporting English Language Learners. The LTG PD intervention targets teachers' mathematics knowledge for teaching geometric similarity, meaning that it is intended to inform both their content and pedagogical content knowledge in ways that foster student learning of this topic.

Results from the Learning and Teaching Geometry PD Field Test

As part of the LTG project, the Foundation Module was field tested in eight sites throughout the United States (Borko, Jacobs, Seago, Mangram, 2014; Seago et al., 2014). Each group was led by a trained facilitator, who was typically a district-level mathematics coordinator or university faculty member. The facilitators each recruited teachers to participate in the professional development (treatment teachers) and a comparison group (colleagues of the treatment teachers). In total, the pilots involved 127 participants: 87 treatment teachers and 40 comparison teachers. The participants included K-12 teachers, teacher leaders, coaches, and pre-service teachers.

The treatment and comparison teachers completed a series of pre- and post-knowledge assessments. The knowledge assessments included a multiple-choice content assessment and two embedded assessments (i.e. incorporated into the PD workshops) each with three open-ended tasks. All of these assessments were used to examine the impact of the professional development on teachers' mathematical knowledge for teaching (MKT) geometric similarity. Two of the field test sites agreed to administer a student assessment to the students of both the treatment (n=162) and comparison teachers (n=104). The student assessment was a multiple-choice test that focused on geometric similarity.

Treatment teachers demonstrated significant knowledge gains relative to the comparison teachers on all of the assessments. Specifically, on the multiple-choice teacher knowledge assessment, the treatment teachers improved their scores by 8.73 percentage points on average. In contrast, the comparison teachers improved by 1.68 percentage points on average. On the embedded assessment tasks (3 tasks per assessment), the treatment teachers showed significant improvement on five of six tasks, whereas the comparison group did not show significant improvement on any task. On the student assessment, students of the treatment teachers on average demonstrated significantly higher gains compared to students of the comparison teachers.

Taken together, the results of the field study showed that the LTG PD positively impacted the participating teachers' knowledge as well as their students' knowledge. These data provide evidence of the promise of the Foundation Module for achieving the intended teacher and student knowledge outcomes. Although persuasive, the empirical evidence supporting the intervention has thus far been based on data in settings where the materials were used within the design research process of the development project. The finalized version of the Foundation Module was not tested, nor was the complete package of the LTG PD materials (i.e., the Foundational Module together with the Extension Modules) used with any group of teachers. Other limitations include the fact that the samples were based on convenience without random assignment or matching; the pilot groups used different versions of the materials (depending on which version was available at the time) with different facilitators and different teacher populations; and the data collected were relatively limited (i.e., no use of knowledge assessments farther removed from the PD content and no data on classroom practices).

Research Questions

In this paper, we present findings from the LTG Efficacy Study which was designed to explore the effectiveness of the full LTG PD intervention using a group-randomized experimental design. We consider the overall effectiveness of the professional development on critical aspects of teacher learning, and explore geographic location-based differences in teacher and student learning. Our research questions are as follows:

RQ1: To what degree did the LTG PD have an effect on the participating teachers and their students?

a. Did teachers in schools that participated in the LTG PD increase their knowledge of geometry and transformations-based geometry beyond that of teachers in the control schools?
b. Did teachers in schools that participate in the LTG PD improve the quality of their classroom instruction beyond that of teachers in the control schools?
c. Did students taught by teachers in schools that participated in the LTG PD increase their knowledge beyond that of students taught by teachers in the control schools?

RQ2: Did the impact of the LTG PD on teachers' knowledge, classroom instruction, and student knowledge vary across the two geographic locations in which it was implemented?

Method

Sample

Participants in the LTG Efficacy Study were 103 secondary mathematics teachers serving grades 6-12 (47% middle school, 53% high school) from two diverse geographic locations, (49% in Location A, 51% in Location B). One of the locations included teachers from multiple school districts and the other location included teachers from a single large district. Both locations included teachers from a range of grade levels and schools, and who used a variety of curricular resources in their mathematics classrooms.

Two locations were included in the study for several reasons. First, the two locations exhibit unique student populations with a large amount of diversity both within and across geographic regions of the United States. Location A includes teachers from multiple suburban districts surrounding a city in the midwestern US. These districts have a relatively large Hispanic and

English Language Learner population. Location B includes teachers from a single, urban school district in the northeastern region of the US. The student population in this district is largely non-white, with the largest proportion of students identifying as Black.

A second important reason these locations were selected is that district stakeholders in both locations had strong existing collaborative ties with members of the research team, as well as an expressed interest in incorporating the LTG PD as a means of improving their teachers' knowledge and instruction in transformations-based geometry. The assistance of these stakeholders was critical to ensuring that a large enough sample of teachers could be recruited and that the districts would continue to support two years of ongoing PD workshops. Finally, including two geographically distinct locations enabled a large enough sample size to compare the treatment teachers to a comparison group, and to look within the treatment group to detect differences across locations.

Randomization was conducted at the school level, with 32 schools and 49 teachers assigned to the treatment group and 35 schools and 54 teachers assigned to the comparison group. Table 1 provides background information on the sample. No statistically significant differences were found by treatment and control groups for the full sample and within the two locations. Three significant differences were detected by location: teachers in Location A had more years of experience teaching mathematics ($p<.05$), more teachers in Location B had graduate degrees ($p<.001$) and more teachers in Location B were certified in special education ($p<01$).

<u>Professional Development Workshops</u>
The Learning and Teaching Geometry professional development workshops for treatment teachers in both locations began in Summer 2016 and continued throughout the 2016-17 academic year. Nine full days of professional development were offered to teachers in each location, beginning with the five-day Foundation Module followed by four days of Extension Modules. On average, treatment teachers attended 7 days of PD (7.6 for Location A teachers and 6.5 for Location B teachers). In Location A all workshops were held during the school day (with substitutes paid for by the project). Location B teachers likely had a lower attendance rate due to the fact that their academic year workshops were held on weekends per district policy.

Control teachers were offered the opportunity to participate in the same LTG PD workshops (including the Foundation Module and Extension Modules) during the 2017-18 school year, once pre- and post-data collection was completed. The same facilitator led all of the workshops in both locations, after taking part in an extensive facilitator preparation process that included a multi-faceted assessment of fidelity. Based on this fidelity assessment process, facilitator was deemed highly capable of using the PD materials as intended and making appropriate context-based decisions and adaptations (Jacobs, Seago & Koellner, 2017). In all cases, the workshops were held first in Location A and then in Location B for pragmatic reasons, largely due to the timing of the academic years in the two locations.

<u>Measures</u>
Pre and post data collected on the effectiveness of the LTG PD included measures of teacher and student knowledge, videos of classroom instruction, and anonymous teacher reflections on the

professional development. Baseline data for all teachers was collected in Spring 2016 (prior to the summer workshops for the treatment teachers), and post data was collected in Spring 2017.

Teacher Knowledge
Four assessments were used to examine impacts on the teachers' knowledge: one developed by the Center for Research in Mathematics & Science Teacher Development at the University of Louisville, and three created by Horizon Research Inc. The assessment developed by the University of Louisville is part of a series called Diagnostic Science Assessments for Middle School Teachers (DTAMS; Saderholm et al., 2010). For this study we administered the DTAMS geometry/measurement assessment, which serves as a general assessment of teachers' geometry and measurement content and pedagogical content knowledge. The DTAMS assessment is divided into two domains, "knowledge" and "mathematics subject." Items in the "knowledge" domain are: facts, conceptual understanding, reasoning/problem solving, and pedagogical content knowledge. Items in the "mathematics subject" domain are: 2-D geometry, 3-D geometry, transformational/coordinate geometry, and measurement. Scoring of the DTAMS assessment, which includes both closed and open-ended items, was conducted by the DTAMS development team in accordance with their usage policy.

Horizon Research, Inc. developed three pre-post assessments to measure the impact of the LTG PD on teachers' knowledge of geometric similarity: a multiple choice assessment and two embedded assessments (Seago et al., 2014). These were the same measures used in the field study described earlier. They were developed in tandem with the creation of the LTG PD materials and are the most proximal measures of teacher learning from the intervention. The Horizon multiple-choice assessment covers five content areas: dilation, properties of similarity, ratio and proportion, scaling, and congruence transformations (i.e., translation, rotation, and reflection). The items were compiled and modified from existing state, national, and international assessment sources. About half of the items are purely content-based; the rest are set in contexts that situate them in the work of teaching geometry. The items were validated by the LTG developers as accurate and appropriate to the content emphasis of the intervention.

Horizon Research Inc.'s embedded assessments consist of a mathematics task and a videocase analysis task. These tasks exist within the Foundation Module and can either be administered as part of the LTG PD (i.e., early in the PD and again later in the PD) or separately (i.e., to the comparison teachers). The mathematics task includes three open-ended items, which focus primarily on the properties of similarity. The videocase analysis task also includes three open-ended items, which focus primarily on dilation. Horizon developed initial scoring rubrics for the embedded assessments; a modified version of the rubrics was applied by our research team after achieving initial inter-rater agreement between 80-93% on the items within the mathematics task and 85-97% on the items within the videocase analysis task.

Classroom Observations
Teachers' mathematics lessons were videotaped and rated using the Math in Common teacher observation protocol (Perry et al., 2015). The Math in Common protocol, developed as part of WestEd's Math in Common study of K12 mathematics instruction, incorporates eight items that capture various elements of lesson quality. The protocol is adapted from the Mathematical

Quality of Instruction (Hill, 2014) and Teaching for Robust Understanding (Schoenfeld, 2013; Schoenfeld et al., 2014) coding instruments.

Items from the MQI address two broad constructs: richness of the mathematics and student engagement in practices recommended by the Common Core State Standards. Each item is rated on a 4-point scale (1=not present, 2=low, 3=middle, 4=high) and applied at the lesson level:
(1) linking between representations (richness of mathematics)
(2) multiple procedures or solution methods (richness of mathematics)
(3) mathematical sense making (richness of mathematics)
(4) student explanations (Common Core student practices)
(5) mathematical reasoning (Common Core student practices)

Items from TRU focus on a single construct: the mathematical content in the lesson. Each item is rated on a 3-point scale (1=novice, 2=apprentice, 3=expert) and applied at the lesson level:
(1) attention to accurate and well-justified mathematics
(2) access to mathematical content
(3) agency, authority and identity

Members of the research team established initial inter-rater agreement of 92.5% on the Math in Common protocol overall, and between 80-100% on each of the items. To ensure that they were applying the items consistently throughout the coding process, they also established mid-point reliability approximately half-way through the set of videotaped lessons. Midpoint inter-rater reliability was 91% on the protocol overall and between 80-100% on each of the items.

Student Knowledge
In addition to the teacher knowledge assessments described earlier, Horizon Research Inc. also developed an assessment to measure the impact of the LTG PD intervention on student knowledge. This multiple-choice assessment was constructed based on cognitive interviews with students to ensure the items were clear, plausible, and had content validity. The assessment targets students' knowledge of transformations-based geometry in five content areas: dilation, properties of similarity, ratio and proportion, scaling, and congruence transformations. These are the same five content areas targeted by the Horizon multiple choice teacher knowledge assessment.

Teacher Reflections
Treatment teachers provided daily reflections on their LTG PD workshop experiences. Teachers responded to open-ended prompts in an anonymous online survey at the end of each workshop. Although the prompts varied slightly from day to day, teachers typically reported on what they learned, what they struggled with, what they hoped to take back to their classrooms, and what feedback they had for the facilitator. Their responses were coded by two members of the research team for tone (positive, negative, in-between) and the nature of their self-assessed learning (content, pedagogy, both, neither). Inter-rater agreement of at least 80% was established for both codes.

<u>Analysis</u>

Repeated measures ANOVA analyses were used to detect patterns of change over time in the treatment teachers' content knowledge, relative to the comparison group, for all four assessments of teacher knowledge, and for the measures of teachers' quality of classroom instruction. Additional analyses of variance explored change in teachers' knowledge and instruction within the two locations.

Hierarachical linear models (HLM 7.0, Raudenbush, Bryk, Cheong, Fai, Congdon, & Du Toit, 2011) were used to estimate the impact of teacher participation in the LTG PD program on student gains in geometry knowledge. Although a three-level hierarchical model can account for the nested data structure of students in classrooms with the same teacher and the teachers within schools. the sparse number of teachers per school (56 teachers in 41 schools) makes it difficult to reliably estimate the variance between teachers within schools. We therefore limited analyses to a two-level hierarchical model with a student level and a classroom level. For the same reason, the analyses of teacher outcomes were limited to the teacher or classroom level only.

Results

Impacts on Teacher Knowledge

Overall changes in knowledge: treatment teachers versus comparison teachers
Significant differences in teacher knowledge pre and post-scores from all four assessments are shown in Table 2. On the most general, distal assessment of geometry knowledge - the DTAMS geometry/measurement assessment - there were no differences between the treatment and comparison groups on the test as a whole or on any of the eight content area subdomains. Time main effects were found for two of the subdomains (reasoning/problem solving and measurement), indicating that the treatment and comparison teachers gained a significant amount knowledge in these areas to about the same degree over time.

Both LTG PD invention and time main effects were found based on Horizon Research, Inc's teacher knowledge assessments, which were more proximal to the content focus of the workshops. On Horizon's multiple-choice assessment, although there was not a statistically significant difference between the groups on the test as a whole, the treatment teachers showed greater improvement relative to the comparison teachers in the area of congruence transformations. On both of Horizon's embedded assessments there were significant interaction effects. The treatment group made significantly greater knowledge gains relative to the comparison group on dilations, properties of similarity, and total geometry.

Taken together, the results from the three Horizon assessments suggest that teachers who participated in the LTG PD gained knowledge in several of the focal content areas, especially congruence transformations and dilation. It is important to note that these findings are not particularly robust across measures, suggesting that the teachers' knowledge gains may not be generalizable across mathematical contexts.

Changes in knowledge by context: Location A teachers versus Location B teachers
Examining the impact on teacher knowledge by location reveals some interesting differences within the treatment group (see Tables 3 and 4). In Location A, treatment teachers made

significant gains in the content area of measurement (DTAMS), dilation (embedded videocase analysis) and total geometry (embedded math task). The treatment gains for both embedded assessments were significantly larger than the control gains.

On the other hand, Location B teachers made significant gains in reasoning/problem solving (DTAMS), congruence transformations (Horizon), and dilation (embedded videocase analysis). Comparison teachers in Location B made no significant gains.

On the whole, teacher knowledge impacts appear to be limited in scope and generalizability, but teachers in the two locations showed some differences in the nature of their learning of the focal mathematics content from the PD. The only area of overlap across the two groups was a gain in knowledge about dilations, as measured by the embedded videocase assessment. Otherwise, teachers in the two locations appeared to learn about particular mathematical concepts (including measurement, reasoning/problem solving, congruence transformations, and properties of similarity) to different degrees, despite the fact that the groups were led by the same facilitator who adhered to highly specified workshop agendas.

Impacts on Classroom Practice

Overall changes in instruction: treatment teachers versus comparison teachers
Compared to the impacts on teacher knowledge, impacts on classroom practice were considerably more robust and pervasive. Table 5 presents the ratings on each of eight categories within three broader instructional practice constructs: richness of the mathematics, engaging students in mathematical practices recommended by the Common Core State Standards, and mathematics content. The treatment group made significant gains over time in both richness of the mathematics and engaging students in mathematical practices. The gains made by the treatment teachers were larger than those made by the comparison teachers for engaging students in mathematical practices and mathematics content.

It is interesting to note that the comparison teachers started off with significantly higher scores on all three broad constructs in their pre-observations relative to the treatment teachers. However over time the comparison teachers' instructional practices generally remained stagnant while the treatment teachers "caught up" to their colleagues as evidenced by the fact that there were no significant differences on their posttest averages.

Looking within these three constructs, paired difference tests indicate significant change for the treatment teachers in three of the individual categories: sense-making ($p<.05$), mathematical reasoning ($p<.01$), and attention to accurate and well-justified mathematical content ($p<.01$). These three individual categories span the broader constructs, indicating that the treatment teachers improved their instruction across a variety of critical areas. Comparison teachers made no changes any of the individual categories.

Changes in instruction by context: teachers in Location A versus Location B
Dramatic differences emerged in the degree and nature of instructional quality improvement by geographic location (see Table 6). Treatment teachers in Location A made significant improvement in the broad construct of engaging students in mathematical practices, and their

gain in this construct was significantly larger than the comparison teachers. Within the individual categories, paired difference tests indicate that in Location A the treatment teachers significantly improved in supporting students' mathematical reasoning.

Treatment teachers in Location B significantly improved in all three broad constructs as well as four of the eight individual categories: supporting reasoning ($p<.01$), promoting sense-making ($p<.01$), student explanations ($p<.05$), and attention to accurate and well-justified mathematical content ($p<.01$).

It is important to temper these findings with the fact that comparison teachers in both locations also made gains. For example, in Location A the comparison group showed significant improvement in providing access to mathematical content ($p<.01$). In Location B the comparison group showed significant improvement in one of the the broad constructs, richness of the mathematics. It seems likely that a variety of efforts were in place in both locations to support teachers to change their instructional practice in addition to the LTG PD intervention. However it is clear that the LTG PD had a positive and widespread effect on classroom teaching, particularly in Location B.

<u>Impacts on Student Knowledge</u>

Changes in student knowledge across groups and locations
Classroom averages from the Horizon student assessment were used to measure the impact of the LTG PD intervention on student knowledge, by comparing gains across the treatment and comparison teacher groups. Table 7 shows the average pre- and posttest scores and gains students made within treatment and control groups. Students in both groups gained in knowledge within all of the content domains except proportion. Importantly, students of the treatment teachers had significantly higher score increases than students of the comparison teachers in the domain of congruence transformations.

As was the case for teacher gains in both content knowledge and instructional practice, student knowledge gains differed by location (see Table 8). Students of treatment teachers in Location A showed significant improvement in all but one of the measured content areas (dilation). However students of comparison teachers in Location A also improved in most content areas, with the exception of proportion. Likewise, students of treatment teachers in Location B gained in all areas except proportion, while students of comparison teachers in Location B also improved all areas except dilation and proportion. In Location B, students of the treatment teachers had a significantly larger gain in their knowledge of congruence transformations than students of the control teachers. Given the large sample size and the one-dimensional nature of the student data, we turn to the multi-level analyses to more accurately estimate LTG PD effects on student knowledge.

Results from HLM analyses
Because of the large degree of improvement shown by the students of the comparison teachers in both locations, it is important to look more closely at the extent to which the LTG PD may have supported student knowledge gains. To explore the degree to which the LTG PD had an impact

on student knowledge, and the specific mechanism through which such an impact may have occurred, we used hierarchical linear modeling (HLM) as an additional analytic tool.

Primarily we were interested in the effect of the LTG PD on student gains in total mathematics knowledge and the five content areas. The classroom-level predictors were the effect of the LTG PD intervention, teacher content knowledge post PD, and the quality of classroom instruction post PD. Two teacher background covariates were also included: current grade level teaching and years teaching in general. Level-1 includes 758 students; Level-2 includes 56 teachers.

Preliminary analysis of the initial unconditional mean model confirmed that there was significant variation among classrooms available to be explained by the model ($ICC=.11$, $p<.001$). Additional preliminary analyses confirmed that neither of the teacher covariates (grade level or years teaching) was a significant predictor of student gains and both were omitted from the final model.

In the conditional model, we first estimated the direct effect of participation in the LTG PD program on average student gain in mathematical knowledge. We found a significant ($p<.05$) and positive effect of LTG PD on student gains in knowledge of congruence transformations. LTG PD explained a significant proportion of variance in gains in transformations knowledge (10.0%, $ES=0.23$, $p=.048$). No significant effects were found for the remaining outcome measures.

Next, location was added to the model but was not found to be a significant predictor. Teacher content knowledge and the three broad constructs of instructional quality were then entered into the model at the classroom level. In the final model shown in Table 9, LTG PD intervention significantly predicted greater student knowledge gains in the domain of congruence transformations ($ES=.23$, $p=.028$) along with instruction that engaged students in mathematical practices recommended by the Common Core State Standards. Together, they explained 27% of the variance in student gains in knowledge of transformations.

Overall, the HLM models provide evidence that taking part in the LTG PD, together with using instructional practices recommended by the standards, supports student knowledge gains in one of the targeted mathematical domains. It is equally important to highlight the non-significant predictors in these models: teacher content knowledge and location. Although the data indicate some improvements in teacher knowledge for those teachers who took part in the intervention, these improvements were not particularly robust and they did not lead to subsequent student knowledge gains when considered simultaneously with participation in the LTG PD program and the quality of classroom instruction. Neither was geographic location a predictor, meaning that while location differences existed, they were not strong enough to predict student gains in a targeted content area when included in a model alongside PD participation and higher quality of instruction in engaging students. It is also important to notice that these student knowledge gains are limited in scope to congruence transformations, and do not apply to all of the mathematical domains targeted by the LTG PD.

Conclusion

LTG PD Intervention Supports Critical Gains

Our first research question is about the effect of the LTG PD on treatment teachers and their students. In particular, we are interested in looking at the impacts of teachers' knowledge, classroom practices, and student knowledge. Based on the results reported in this paper, there were critical gains in all of these areas, with the largest gains being found in the participating teachers' instructional practice.

The improvements in teacher knowledge were mostly picked up by highly proximal embedded assessments, with very few significant increases occurring on the multiple-choice Horizon assessment or the more distal DTAMS assessment. Furthermore, the treatment teachers' knowledge improvements were confined to only a few of the targeted mathematical domains, especially congruence transformations (the rigid transformations: translation, rotation, and reflection), dilations (a non-rigid transformation), and measurement. Although these topics were a central focus of the PD, it is somewhat disappointing that the knowledge gains were not more robust across measures and more widespread across the targeted content areas.

At the same time, it is encouraging that the treatment teachers made large instructional improvements despite their comparatively smaller gains in content knowledge. These improvements in classroom practice were quite robust and held true across a range of constructs. Importantly, the teachers were filmed teaching any mathematical topic of their choosing, meaning that the improvements are not limited to a single content area or only the geometric content covered by the LTG PD. These findings suggest that it is possible for PD to support mathematics teachers' to make widespread instructional changes in the desired direction. Further, these changes occurred despite the lack of strong impacts on their mathematical knowledge, and despite the fact that the content focus of the PD was relatively narrow.

As we know from the mixed literature on the effectiveness of PD efforts discussed in the introduction, generating improvements in classroom instruction is no small feat for a given intervention, and it is arguably a more important goal than improving teacher knowledge. It is interesting to consider whether an improvement trajectory that ends in increased student learning requires first an increase in teacher knowledge or whether improved classroom practice by itself is sufficient. We do not contend that our study answers this question, but it does raise the issue and further investigation as to the relationship between teacher knowledge gains, classroom practice improvements and student learning is certainly warranted. It is also possible that the nature our results (strong impacts on instructional practice and much smaller impacts on teacher knowledge) have to do the with particular nature of our PD program – which included a heavy focus on the collaborative analysis of classroom video. Again, this is a question that could be investigated by future research.

Participating in the LTG PD and making subsequent instructional improvements (especially engaging students in recommended mathematical practices) both were directly related to student learning in the area of congruence transformations, as suggested by the HLM analyses. Congruence transformations was only one of several content areas targeted by the LTG PD, and it is interesting to consider why this area in particular might have demonstrated the largest increase for the students of the treatment teachers. One possibility is that the treatment teachers themselves gained knowledge in this area, and the topic is widely relevant across grades and mathematical content areas. Although we did not collect data on teachers' instruction of transformations, we can speculate that it may have been a topic more widely and deeply covered by the treatment teachers based on their experiences in the workshops. Our data show that the comparison students also gained substantially in all of the targeted content areas, suggesting that mathematics teachers as a group are attending to these mathematical topics in their instruction and appear effective in supporting their learning. However, on the Horizon assessment students generally scored lowest in the areas of ratio, dilation and congruence transformations, pointing to a need for continued attention to their learning of these topics.

Differences Across the Two Locations

Most of the literature on the design and implementation of effective professional development cautions that context is a critical factor to keep in mind (e.g. Cohen & Ball, 1999; Loucks-Horsley et al., 2003). As discussed in the introduction, professional learning by design takes place within a given social and policy-based context, including at the school, district, state and national levels. Additionally, a good deal of professional learning (including in the LTG intervention) takes place at the community level, amongst a given group of teachers who meet face-to-face at regular intervals throughout an academic year. Given all of these contextual factors, it is not surprising that learning outcomes might differ across locations.

Our second research question addresses the differential impacts of participation in the LTG PD based on the geographic locations of the workshops. Including teachers in two distinct areas of the US enables us to explore context-based differences relevant to the effectiveness of the PD in the areas of teacher knowledge, classroom practice, and student learning. We found differences across the locations in each of these three areas -- meaning that the treatment teachers gained different types of content knowledge, improved on different aspects of their classroom instruction, and had students who learned about somewhat different aspects of geometric similarity depending on their location. It is important to note that these "differences" are based on statistical analyses within the treatment sample, looking over time. In some cases the treatment groups' improvements were significantly greater than the comparison teachers in the same location, and in other cases they were not.

Our aim is not to defend a position that the two groups of treatment teachers were indeed different in their learning outcomes, but rather to highlight the possibility that such differences could exist and to consider why this might be the case. One reason for a differential impact across teacher groups might be that they participated in and experienced the PD differently. We have some evidence that the groups did in fact have very different impressions of the LTG PD workshops, despite the fact that they were conducted by the same facilitator using the same resources. In their anonymous reflections submitted at the conclusion of each workshop, Location A teachers were markedly more critical of their experiences than Location B teachers (see Table 10). Less than half of the teachers (41%) in Location A described the Foundation Module (the first 5 workshops) as a positive experience overall. By contrast, approximately two thirds of the teachers (67%) in Location B reported that they had a positive experience. Conversely, 28% of the Location A teachers indicated that the LTG PD was a negative experience, compared to only 2% of the Location B teachers.

There are many possible explanations for the differences across locations, all of which are worth further exploration. For example, the teachers in Location were more experienced teachers in general and in teaching math. Several expressed a senses of disappointment by the perceived lack of challenge related to the mathematical content. Also because the Location B workshops were conducted after the Location A workshops, the facilitator might have made subtle changes that resulted in more productive and in-depth conversations. Another possibility is that the community was more tightly formed in Location B, which led to stronger engagement and interest in the material. We plan to look carefully at the nature of the discussions and participation structures within the workshops in both locations, in order to better understand how the teachers communicated with one another and whether there were differences across the groups.

It is important to point out that although there appear to be relevant differences across geographic locations, it is not the case that the LTG PD impacts were limited to teachers in a given group. For example, in both locations there was a substantial impact on teachers' classroom instruction, although the impact appears to be more robust for teachers in Location B given that they improved on all three broad practice constructs whereas the teachers in Location A only improved on one of the three broad constructs. The construct that teachers in Location A improved on, student engagement in mathematical practices, was found to predict student knowledge gains. Some important conclusions from these findings are that (1) teachers can be relatively negative in their perceptions of the PD but still show significant gains in meaningful areas, and (2) teachers' collaborative experience of the PD might translate into different types of impacts across groups.

Implications

This efficacy study of the Learning and Teaching Geometry professional development intervention has implications for research on the design and implementation of efforts to support teacher learning. Our findings suggest that a well-specified, video-based intervention can be implemented to improve critical aspects of teacher knowledge and (perhaps more importantly) instructional practice, leading to gains in student learning in a targeted mathematical content area. Additionally this study points to the importance of using measures that are capable of detecting critical gains, especially with respect to instructional practice, and of looking within treatment teacher groups across locations to identify relevant context-based differences that may affect their experiences and learning.

References

Academy for Educational Development (2002). *National Writing Project: Final evaluation report*. Author: New York, NY.

Akkerman, S. F., & Bakker, A. (2011). Learning at the boundary: An introduction. *International Journal of Educational Research*, *50*(1), 1-5.

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

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., ... & Tsai, Y. M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American educational research journal*, *47*(1), 133-180.

Bell, C. A., Wilson, S. M., Higgins, T., & McCoach, D. B. (2010). Measuring the effects of professional development on teacher knowledge: The case of developing mathematical ideas. *Journal for Research in Mathematics Education*, 479-512.

Blomberg, G., Renkl, A., Sherin, M. G., Borko, H., & Seidel, T. (2013). Five research-based heuristics for using video in pre-service teacher education. *Journal for Educational Research Online/Journal Für Bildungsforschung Online*, *5*(1), 90-114.

Blomberg, G., Sherin, M. G., Renkl, A., Glogger, I., & Seidel, T. (2014). Understanding video as a tool for teacher education: Investigating instructional strategies to promote reflection. *Instructional Science*, *42*(3), 443-463.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn* (Vol. 11). Washington, DC: National Academy Press.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational researcher*, *18*(1), 32-42.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational researcher*, *33*(8), 3-15.

Borko, H., Jacobs, J., Eiteljorg, E., & Pittman, M.E. (2008). Video as a tool for fostering productive discourse in mathematics professional development. *Teaching and Teacher Education*, *24*, 417-436.

Borko, H., Jacobs, J., & Koellner, K. (2010). Contemporary approaches to teacher professional development. P. Peterson, E. Baker & B. McGaw (Eds.), *International Encyclopedia of Education, Vol 7* (pp. 548-556). Oxford: Elsevier.

Borko, H., Jacobs, J., Seago, N. & Mangram, C. (2014). Facilitating video-based professional development: Planning and orchestrating productive discussions. In Y. Li, E.A. Silver & S. Li

(Eds.), *Transforming mathematics instruction: Multiple approaches and practices* (pp. 259-281).Dordrecht, Netherlands: Springer International Publishing Switzerland.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational researcher*, *18*(1), 32-42.

Cobb, P., McClain, K., de Silva Lamberg, T., & Dean, C. (2003). Situating teachers' instructional practices in the institutional setting of the school and district. *Educational Researcher*, *32*(6), 13-24.

Cohen, D. K., & Ball, D. L. (1999). Instruction, capacity, and improvement. Consortium for Policy Research in Education, University of Pennsylvania. Retrieved from: https://files.eric.ed.gov/fulltext/ED431749.pdf

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.

Firestone, W. A., Mangin, M. M., Martinez, M. C., & Polovsky, T. (2005). Leading coherent professional development: A comparison of three districts. *Educational Administration Quarterly*, *41*(3), 413-448.

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.

Gallagher, H. A., Woodworth, K. R., & Arshan, N. L. (2015). *Impact of the national writing project's college-ready writers program on teachers and students*. Menlo Park, CA: SRI International.

Gaudin, C., & Chaliès, S. (2015). Video viewing in teacher education and professional development: A literature review. *Educational Research Review*, *16*, 41-67.

Gewertz, C. (2013). Teachers say they are unprepared for Common Core. *Education Week*, *32*(22), 1-12.

Greeno, J. G., Collins, A. M., & Resnick, L. B. (1996). Cognition and learning. *Handbook of educational psychology*, *77*, 15-46.

Grissmer, D. W., Flanagan, A., Kawata, J. H., Williamson, S., & LaTourrette, T. (2000). *Improving student achievement: What state NAEP test scores tell us*. Santa Monica: CA. Rand Corporation.

Grossman, P., Wineburg, S., & Woolworth, S. (2001). Toward a theory of teacher community. *Teachers College Record*, 103(6), 942-1012.

Hatch, T., Shuttleworth, J., Jaffee, A.T. & Marri, A. (2016). Videos, pairs, and peers: What connects theory and practice in teacher education? *Teaching and Teacher Education*, *59*, 274-284.

Hill, H. C. (2014). *Mathematical Quality of Instruction (MQI): 4-point version*. Ann Arbor, MI: University of Michigan Learning Mathematics for Teaching Project.

Ingvarson, L., Meiers, M., & Beavis, A. (2005). Factors affecting the impact of professional development programs on teachers' knowledge, practice, student outcomes & efficacy. *Education Policy Analysis Archives, 13(10).*

Jacob, R., Hill, H., & Corey, D. (2017). The impact of a professional development program on teachers' mathematical knowledge for teaching, instruction, and student achievement. *Journal of Research on Educational Effectiveness*, *10*(2), 379-407.

Jacobs, J., Seago, N. & Koellner, K. (2017). Preparing facilitators to use and adapt professional development materials productively. *International Journal of STEM Education*, *4*(1), 30.

Kazemi, E., & Hubbard, A. (2008). New directions for the design and study of professional development: Attending to the coevolution of teachers' participation across contexts. *Journal of teacher education*, *59*(5), 428-441.

Kennedy, M. M. (2016). How does professional development improve teaching?. *Review of Educational Research*, *86*(4), 945-980.

Kersting, N. (2008). Using video clips of mathematics classroom instruction as item prompts to measure teachers' knowledge of teaching mathematics. *Educational and Psychological Measurement*, *68*(5), 845-861.

Kersting, N. B., Givvin, K. B., Sotelo, F. L., & Stigler, J. W. (2010). Teachers’ analyses of classroom video predict student learning of mathematics: Further explorations of a novel measure of teacher knowledge. *Journal of Teacher Education*, *61*(1-2), 172-181.

Kersting, N. B., Givvin, K. B., Thompson, B. J., Santagata, R., & Stigler, J. W. (2012). Measuring usable knowledge: Teachers’ analyses of mathematics classroom videos predict teaching quality and student learning. *American Educational Research Journal*, *49*(3), 568-589.

Koellner, K. & Jacobs, J. (2015). Distinguishing models of professional development: The case of an adaptive model’s impact on teachers’ knowledge, instruction, and student achievement. *Journal of Teacher Education, 66*(1), 51-67.

Koh, K. (2015). The use of video technology in pre-service teacher education and in-service teacher professional development. In *Cases of mathematics professional development in east asian countries* (pp. 229-247). Springer, Singapore.

Kutaka, T. S., Smith, W. M., Albano, A. D., Edwards, C. P., Ren, L., Beattie, H. L., ... & Stroup, W. W. (2017). Connecting teacher professional development and student mathematics achievement: a 4-year study of an elementary mathematics specialist program. *Journal of Teacher Education*, *68*(2), 140-154.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.

Little, J. W. (2003). Inside teacher community: Representations of classroom practice. *Teachers College Record*, *105*(6), 913-945.

Loucks-Horsley, S., Stiles, K. E., Mundry, S., Love, N., & Hewson, P. W. (2009). *Designing professional development for teachers of science and mathematics*. Thousand Oaks: CA. Corwin Press.

Major, L., & Watson, S. (2018). Using video to support in-service teacher professional development: the state of the field, limitations and possibilities. *Technology, Pedagogy and Education*, *27*(1), 49-68.

McCallum, W. (2011). Gearing up for the Common Core State Standards in mathematics. *Draft report of the policy workshop.* Available from: http://internet.math.arizona.edu/~ime/2010-11/2011_04_01_IME_PD.pdf

Miller, K., & Zhou, X. (2007). Learning from classroom video: What makes it compelling and what makes it hard. In Goldman-Segal, R. & Pea, R. (Eds.), *Video Research in the Learning Sciences*, Hillsdale, NJ: Lawrence Earlbaum, 321-334.

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.

Perry, R. R., Finkelstein, N. D., Seago, N., Heredia, A., Sobolew-Shubin, S., & Carroll, C. (2015). *Taking stock of Common Core math implementation: Supporting teachers to shift instruction*. San Francisco, CA: WestEd.

Perry, R., Seago, N., Burr, E., Broek, M., Finkelstein, N. (2015). *Classroom observations: Documenting shifts in instruction for districtwide improvement.* San Francisco, CA: WestEd.

Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning?. *Educational researcher*, *29*(1), 4-15.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R., & Du Toit, M. (2011). *Hierarchical linear and nonlinear modeling (HLM7)*. Lincolnwood, IL: Scientific Software International.

Roth McDuffie, A., Drake, C., Choppin, J., Davis, J. D., Magaña, M. V., & Carson, C. (2017). Middle school mathematics teachers' perceptions of the Common Core State Standards for Mathematics and related assessment and teacher evaluation systems. *Educational Policy*, *31*(2), 139-179.

Saderholm, J., Ronau, R., Brown, E. T., & Collins, G. (2010). Validation of the diagnostic teacher assessment of mathematics and science (DTAMS) instrument. *School Science and Mathematics*, *110*(4), 180-192.

Santagata, R., Kersting, N., Givvin, K. B., & Stigler, J. W. (2010). 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.

Schoenfeld, A. H. (2013). Classroom observations in theory and practice. *ZDM*, *45*(4), 607-621.

Schoenfeld, A. H., & Floden, R. E. & The Algebra Teaching Study and Mathematics Assessment Project. (2014). *The TRU Math Scoring Rubric*. Berkeley, CA & E. Lansing, MI: Graduate School of Education. University of California, Berkeley & College of Education, Michigan State University. Retrieved from http://ats. berkeley. edu/tools. html.

Scribner, J. P. (1999). Professional development: Untangling the influence of work context on teacher learning. *Educational Administration Quarterly*, *35*(2), 238-266.

Seago, N., Jacobs, J., Driscoll, M., Callahan, P., Matassa, M. & Nikula, J. (2017). *Learning and teaching geometry: Video cases for mathematics professional development, grades 5-10*. San Francisco, CA: WestEd.

Seago, N., Jacobs, J., Driscoll, M., Nikula, J., Matassa, M. & Callahan, P. (2013). Developing teachers' knowledge of a transformations-based approach to geometric similarity. *Mathematics Teacher Educator*, *2*(1), 74-85.

Seago, N.M., Jacobs, J.K., Heck, D.J., Nelson, C.L. & Malzahn, K.A. (2014). Impacting teachers' understanding of geometric similarity: Results from field testing of the Learning and Teaching Geometry professional development materials. *Professional Development in Education*, 40(4), 627-653.

Sherin, M. G. (2007). The development of teachers' professional vision in video clubs. *Video research in the learning sciences*, 383-395.

Sherin, M., Jacobs, V., & Philipp, R. (Eds.). (2011). *Mathematics teacher noticing: Seeing through teachers' eyes*. New York, NY: Routledge.

Sherin, M., & Van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *Journal of teacher education*, *60*(1), 20-37.

Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, *19*(3), 387-420.

Sztajn, P., Marrongelle, K. A., Smith, P., & Melton, B. L. (2012). Supporting implementation of the Common Core State Standards for Mathematics: Recommendations for professional development. Raleigh, NC: The William and Ida Friday Institute for Educational Innovation at the North Carolina State University College of Education.

Taylor, J. A., Roth, K., Wilson, C. D., Stuhlsatz, M. A., & Tipton, E. (2017). The effect of an analysis-of-practice, videocase-based, teacher professional development program on elementary students' science achievement. *Journal of Research on Educational Effectiveness*, *10*(2), 241-271.

Teuscher, D., Tran, D., & Reys, B. J. (2015). Common Core State Standards in the middle grades: What's new in the geometry domain and how can teachers support student learning?. *School Science and Mathematics*, *115*(1), 4-13.

Tran, D., Reys, B. J., Teuscher, D., Dingman, S., & Kasmer, L. (2016). Analysis of curriculum standards: An important research area. *Journal for Research in Mathematics Education*, *47*(2), 118-133.

van Es, E. A. (2012). Using video to collaborate around problems of practice. *Teacher Education Quarterly*, *39*(2), 103-116.

van Es, E. A., & Sherin, M. G. (2008). Mathematics teachers'"learning to notice" in the context of a video club. *Teaching and teacher education*, *24*(2), 244-276.

van Es, E. A., & Sherin, M. G. (2010). The influence of video clubs on teachers' thinking and practice. *Journal of Mathematics Teacher Education*, *13*(2), 155-176.

Wayne, A. J., Yoon, K. S., Zhu, P., Cronen, S., & Garet, M. S. (2008). Experimenting with teacher professional development: Motives and methods. *Educational researcher*, *37*(8), 469-479.

Webster-Wright, A. (2009). Reframing professional development through understanding authentic professional learning. *Review of educational research*, *79*(2), 702-739.

Wei, R. C., Darling-Hammond, L., & Adamson, F. (2010). *Professional development in the United States: Trends and challenges* (Vol. 28). Dallas, TX: National Staff Development Council.

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). *Reviewing the evidence on how teacher PD affects student achievement*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest.

Table 1

Teacher Background Survey Responses, by Group and Location – percent of teachers

Background	All Teachers (N=103)	Location A (n=50)	Location B (n=53)
Current grade level			
Middle school	47	37	40
High school	53	63	60
Grade levels taught			
Elementary	21	30	14
Middle school	64	64	64
High school	80	84	76
Years teaching			
0-5	40	32	48
6-10	28	22	35
11-20	25	36	13
>20	7	10	4
Years teaching math			
0-5	45	36	53
6-10	27	24	30
11-20	22	32	13
>20	6	8	4
Teaching certification			
Elementary education	19	26	13
Elementary math	2	4	0
Middle math	45	42	47
Secondary math	80	82	77
Undergraduate or graduate degree in Math, Math Education	79	64	93

Note. No differences detected on background measures between Treatment and Control teachers either across the full sample or within location. Percent may be greater than 100 because some participants responded in more than one category.

Table 2

Means and Standard Deviations of Average Teacher Scores on Geometry Measures over Time (N=103 Teachers)

	Pre PD		Post PD		Change over Time	
	Treatment (*n*=49)	Control (*n*=54)	Treatment (*n*=49)	Control (*n*=54)	Treatment (*n*=49)	Control (*n*=54)
Subject	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M*	*M*
DTAMS						
Reasoning, problem solving	6.52 (2.25)	5.98 (2.27)	7.25 (2.43)	6.71 (2.53)	.73**	.73**
Measurement	5.38 (2.00)	5.20 (2.17)	6.40 (1.95)	5.96 (2.14)	1.02**	.76*
Horizon						
Cogruence Transformations	3.04 (1.03)	2.74 (1.13)	3.43 (.91)	2.81 (1.06)	.39**	.06
Embedded videocase task						
Dilations[1]	14.13 (5.10)	13.26 (6.04)	17.04 (3.04)	13.13 (6.03)	2.91***	-.13
Properties of Similarity[1]	3.09 (1.74)	2.90 (2.00)	3.65 (1.89)	2.58 (2.02)	.56*	-.33
Embedded math task						
Total Geometry[1]	9.02 (2.86)	8.31 (3.01)	10.47 (3.34)	7.48 (3.23)	1.45**	-.83

Note. Significant change over time indicated by *$p<.05$; **$p<.01$; ***$p<.001$.

[1]Treatment gain significantly larger than Control gain for Dilations ($p<.01$, $ES=.54$), Properties of Similarity ($p<.05$, $ES=.49$), and Total Geometry ($p<.01$, $ES=.60$).

Table 3

Means and Standard Deviations of Average Teacher Scores on Geometry Measures over Time within Location A (n=50 Teachers)

	Pre PD		Post PD		Change over Time	
	Treatment (n=25)	Control (n=25)	Treatment (n=25)	Control (n=25)	Treatment (n=25)	Control (n=25)
Subject	M (SD)	M (SD)	M (SD)	M (SD)	M	M
DTAMS						
Reasoning, problem solving	6.56 (1.96)	6.28 (2.46)	7.12 (2.30)	7.32 (2.29)	.56	1.08**
Measurement	5.52 (2.06)	5.72 (2.13)	6.56 (1.81)	6.60 (1.66)	1.04*	.88
Horizon						
Congruence Transformations	3.42 (.78)	2.88 (1.15)	3.63 (.71)	3.00 (1.18)	.21	.13
Embedded videocase task						
Dilations[1]	15.17 (4.53)	15.28 (4.86)	17.96 (1.81)	14.36 (5.14)	2.79**	-.92
Properties of Similarity	3.50 (1.56)	3.40 (1.76)	4.08 (1.82)	3.12 (1.88)	.58	-.28
Embedded math task						
Total Geometry[1]	9.40 (2.93)	7.96 (3.26)	11.72 (1.90)	8.04 (2.48)	2.32***	.08

Note. Significant change over time indicated by *p<.05; **p<.01; ***p<.001.
[1]Treatment gain significantly larger than Control gain for Dilations (p<.01, ES=.78) and Total Geometry (p<.01, ES=.77).

Table 4

Means and Standard Deviations of Average Teacher Scores on Geometry Measures over Time within Context Location B (n=47 Teachers)

	Pre PD		Post PD		Change over Time	
	Treatment (n=24)	Control (n=29)	Treatment (n=24)	Control (n=29)	Treatment (n=24)	Control (n=29)
	M (SD)	M (SD)	M (SD)	M (SD)	M	M
DTAMS						
Reasoning, problem solving	6.48 (2.57)	5.67 (2.06)	7.39 (2.61)	6.08 (2.65)	.91*	.42
Measurement	5.22 (1.95)	4.67 (2.12)	6.22 (2.13)	5.29 (2.40)	1.00	.63
Horizon						
Congruence Transformations	2.64 (1.14)	2.61 (1.12)	3.23 (1.07)	2.61 (.89)	.59**	.00
Embedded videocase task						
Dilations	13.00 (5.54)	11.19 (6.49)	16.05 (3.77)	12.00 (6.65)	3.05*	.59
Proporties of Similarity	2.64 (1.84)	2.44 (2.14)	3.18 (1.89)	2.07 (2.06)	.55	-.37
Embedded math task						
Total Geometry	8.63 (2.80)	8.62 (2.80)	9.17 (4.00)	7.00 (3.74)	.54	-1.62*

Note. Significant change over time indicated by *p<.05; **p<.01. No differences in gains detected between groups.

Table 5
Means and Standard Deviations of Average Teacher Ratings on Observed Quality of Instruction over Time (N=92 Teachers)

	Pre PD		Post PD		Change over Time	
	Treatment (*n*=44)	Control (*n*=48)	Treatment (*n*=44)	Control (*n*=48)	Treatment (*n*=44)	Control (*n*=48)
Rating	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M*	*M*
Richness of mathematics[1]	1.96 (.73)	2.31 (.85)	2.28 (.89)	2.47 (.94)	.33**	.16
Student engagement in mathematical practices[2]	1.82 (.76)	2.39 (1.04)	2.18 (.96)	2.31 (.98)	.36**	-.08
Mathematics (content, ideas, access, agency)[3]	1.87 (.53)	2.23 (.57)	2.05 (.65)	2.09 (.67)	.18	-.14

Note. Significant change over time indicated by $^{**}p<.01$.
[1]Average of Linking, Multiple methods, and Sense making ratings; ratings coded 1=Not Present, 2=Low, 3=Middle, 4=High.
[2]Average of Explanation and Reasoning ratings; ratings coded 1=Not Present, 2=Low, 3=Middle, 4=High. Treatment gain significantly larger than Control gain ($p<.01$, $ES=.54$).
[3]Average of Mathematics, Access, and Agency ratings; ratings coded 1=Novice, 2=Apprentice, 3=Expert. Treatment gain significantly larger than Control gain ($p<.05$, $ES=.45$).

Table 6

Means and Standard Deviations of Average Teacher Ratings on Observed Quality of Instruction over Time, by Context (N=92 Teachers)

	Pre PD		Post PD		Change over Time	
	Treatment (*n*=23)	Control (*n*=25)	Treatment (*n*=23)	Control (*n*=25)	Treatment (*n*=23)	Control (*n*=25)
Rating	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M*	*M*
Location A (*n*=48)						
Richness of mathematics[1]	2.09 (.73)	2.21 (.82)	2.35 (.93)	2.12(.89)	.26	-.09
Student engagement in mathematical practices[3]	1.70 (.67)	2.28 (1.06)	2.02(.94)	2.14 (.97)	.33*	-.14
Mathematics [4]	1.93 (.56)	2.23 (.57)	2.00 (.74)	1.95 (.69)	.07	-.28*
Location B (*n*=44)						
	Pre PD		Post PD		Change over Time	
	Treatment (*n*=21)	Control (*n*=24)	Treatment (*n*=21)	Control (*n*=24)	Treatment (*n*=21)	Control (*n*=24)
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*M*	*M*
Richness of mathematics[1]	1.81 (.72)	2.29 (.91)	2.21 (.85)	2.76 (.80)	.40**	.46*
Student engagement in mathematical practices[2]	1.95 (.84)	2.33(1.06)	2.35(.98)	2.35(.87)	.40*	.01
Mathematics[4]	1.81 (.50)	2.08 (.64)	2.10 (.56)	2.15 (.48)	.29**	.07

Note. Significant change over time indicated by *p<.05; **p<.01.

[1]Average of Linking, Multiple methods, and Sense making ratings; ratings coded 1=Not Present, 2=Low, 3=Middle, 4=High.

[2]Average of Explanation and Reasoning ratings; ratings coded 1=Not Present, 2=Low, 3=Middle, 4=High

[3]Location A Treatment gain significantly larger than Location A Control gain (p<.05,ES=.61).

[4]Average of Mathematics, Access, and Agency ratings; ratings coded 1=Novice, 2=Apprentice, 3=Expert.

Table 7

Means and Standard Deviations of Average Student Gains on Geometry Measures over Time (N=758 Students)

	Pretest		Posttest		Change over Time	
	Treatment (*n*=375)	Control (*n*=383)	Treatment (*n*=375)	Control (*n*=383)	Treatment (*n*=375)	Control (*n*=383)
Horizon Subject	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M*	*M*
Total average	1.94 (.80)	1.63 (.73)	2.26 (.87)	1.95 (.83)	.32***	.31***
Dilation	1.54 (.95)	1.27 (.86)	1.73 (.94)	1.42 (.91)	.19***	.16**
Scaling	2.81 (1.56)	2.29(1.44)	3.30 (1.65)	2.76(1.58)	.51***	.48***
Ratio	2.68 (1.33)	2.16(1.35)	3.03 (1.36)	2.72(1.36)	.39***	.57***
Proportion	1.20 (.74)	1.06 (.72)	1.24 (.70)	1.11 (.69)	.04	.06
Congruence Transformations[1]	1.57 (.99)	1.53 (.91)	2.09 (1.02)	1.80(1.07)	.51***	.28***

Note. Significant change over time indicated by **$p<.01$; ***$p<.001$.
[1]Treatment gain significantly larger than Control gain, $p<.01$, *ES*=.20.

Table 8

Means and Standard Deviations of Average Student Gains on Geometry Measures over Time, by Location (N=758 Students)

	Pretest		Posttest		Change over Time	
	Treatment (*n*=375)	Control (*n*=383)	Treatment (*n*=375)	Control (*n*=383)	Treatment (*n*=375)	Control (*n*=383)
Horizon Subject	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M*	*M*
Location A (*n*=397)						
Total average	1.89 (.80)	1.72 (.74)	2.19 (.90)	2.04 (.85)	.30***	.32***
Dilation	1.51 (.94)	1.19 (.89)	1.56 (.92)	1.46 (.90)	.06	.25**
Scaling	2.69 (1.56)	2.37 (1.52)	3.22 (1.65)	2.81 (1.60)	.56***	.41***
Ratio	2.66 (1.32)	2.38 (1.33)	2.98 (1.34)	2.89 (1.33)	.38***	.51***
Proportion	1.10 (.74)	1.16 (.73)	1.22 (.72)	1.18 (.73)	.12*	.03
Congruence Transformations	1.63 (1.00)	1.57 (.96)	2.01 (1.09)	1.95 (1.16)	.38***	.37***
Location B (*n*=426)						
Total average	2.00 (.79)	1.54 (.71)	2.35 (.84)	1.85 (.80)	.35***	.31***
Dilation[1]	1.57 (.96)	1.34 (.82)	1.92 (.93)	1.38 (.92)	.34***	.07
Scaling	2.96 (1.56)	2.20 (1.35)	3.40 (1.65)	2.72 (1.57)	.44***	.56***
Ratio	2.71 (1.34)	1.93 (1.34)	3.09 (1.38)	2.55 (1.37)	.41***	.64***
Proportion	1.32 (.73)	.96 (.69)	1.26 (.69)	1.04 (.65)	-.06	.09
Congruence Transformations[2]	1.50 (.97)	1.48 (.86)	2.17 (.92)	1.66 (.94)	.67***	.19*

Note. Significant change over time indicated by *p<.05; **p<.01; ***p<.001.

[1]Location B Treatment gain in Dilation knowledge significantly larger than Location B Control gain (p<.05, *ES*=.27).

[2]Location B Treatment gain significantly larger than Location B Control gain (p<.001, *ES*=.43).

Table 9

Results of HLM ANOVA on Student Gain in Knowledge of Transformations

Effects				
Initial Random Effects	Variance component	χ2	*p*-Value	Variance explained
Mean of Transformations Gain, u_{0j}	.141	143.486	<.001	27.2%[1]
Level-1 effect, r_{ij}	1.194			
Final Fixed Effects	Coefficient	*SE*	*p*-Value	
Mean of Transformations Gain, γ_{00}	.372	.061	<.001	
LTG PD effect, γ_{01}	.268[2]	.119	.028	
Quality of instruction – Student engagement, γ_{02}	.166	.067	.017	
Final Random Effects	Variance Component	χ2	*p*-Value	
Mean of Transformations Gain, u_{0j}	.103	116.984	<.001	
Level-1 effect, r_{ij}	1.195			

Note. *N*=758 students, *N*=56 teachers.

[1]Both LTG PD and Quality of instruction explained 27.2% of the variance in Student Transformations Gain. The addition of Quality of Instruction explained 20.2% of the variance beyond the effect of LTG PD.

[2]*ES*=.23. The effect size for hierarchical linear modeling analyses with cluster-level assignment is the adjusted group mean difference divided by the unadjusted pooled within-group standard deviation (Feingold, 2009; Raudenbush & Liu, 2001; What Works Clearinghouse, 2008).

Table 10

Teacher Perception of LTG PD Foundation Module Experience – Percent of teachers (n=41 Treatment teachers)

General tone	**Location A (*n*=24)**	**Location B (*n*=17)**
Positive	41%	67%
Some positive and Some Negative	20%	15%
Negative	28%	2%
Can't tell	10%	16%