Classroom discourse in mathematics lessons: the effect of a hybrid practice-based professional development program

Zehavit Kohen & Hilda Borko

Routledge
Taylor & Francis Group

Check for updates

ARTICLE

Classroom discourse in mathematics lessons: the effect of a hybrid practice-based professional development program

Zehavit Kohen and Hilda Borko

Faculty of Education in Science and Technology, Technion, Israel Institute of Technology

ABSTRACT

There is a need to strengthen the theory-practice connection, particularly for early-career mathematics teachers, through effective professional development (PD) programs that integrate theory and research with the implementation of classroom practices. The current study investigates the effect of the online coaching component a hybrid practice-based-PD program that is focused on supporting teachers' facilitation of classroom discourse through observing and reflecting on their instructional practices during video-recorded classroom lessons. Participants were 18 early-career mathematics teachers who participated in a two-year Fellowship Program. Teachers' discourse instructional practice was analysed quantitatively using the PLATO observation tool. The relationship between these instructional practices and the online component of the PD was examined qualitatively through the analysis of teachers' reflections and coach's feedback during recorded coaching conversations. Findings revealed an improvement in teachers' discourse practices during video-recorded mathematics lessons. This improvement was associated with more explicit, specific references to classroom discourse in the teachers' reflections and more positive feedback from the coach during online coaching conversations. The study makes theoretical, methodological and practical contributions regarding effectiveness of practice-based PD programs focusing on classroom videos and coaching.

Introduction

Recent research has stressed the necessity of strengthening the theory-practice connection during professional development (PD) for mathematics teachers (Desimone and Garet 2015). This focus has considerable importance, particularly for early career teachers. Many of these young teachers tend to leave the educational system during their initial professional experience, as they feel unprepared and unable to apply teaching theory when encountering diverse educational needs in actual practice in the school environment (Liston *et al.* 2006, Ingersoll and Perda 2009, Banks *et al.* 2014).

Researchers attribute the difficulties in connecting theory and practice to a lack of practice-based PD programs that integrate theory and research with practice implementation. They therefore suggest sustained and coherent PD programs that use artefacts such as classroom video (Borko *et al.* 2008, Gomez Zaccarelli *et al.* 2018) and student work (Kazemi and Franke 2004) to situate the PD activities in practice (Ball and Cohen 1999, Borko *et al.* 2010, Gibson and Ross 2016).

Unfortunately, many PD programs for teachers are not designed to provide continuing support for the teachers who are attempting to implement new curricula or pedagogies (Barnett 2002, Dede *et al.* 2009). A major factor underlying the nearly 50% attrition rate among early career teachers, especially within their first 5 years of classroom experience, is the absence of day-to-day

---

professional support such as coaching. Current approaches to PD generally fail to provide such assistance (National Commission on Teaching and America's Future 2003).

Not only do new mathematics teachers face challenges common to all new teachers, but in the US as in other countries, they have the added complexity of connecting mathematical practices to mathematical content (Standards for Mathematical Content: Common Core. 2016). According to President's Council of Advisors on Science and Technology (2010) report, there is a disconnect between what mathematics teachers need to know in order to accomplish desired learning most effectively over the full range of mathematics proficiency, and the reality of instructional practice for most teachers in a real classroom setting. PD programs for teachers should offer opportunities for enhancing both the content and pedagogical knowledge required in a specific context of mathematics teaching (Loucks-Horsley *et al.* 2010) – teaching that will foster students' conceptual understanding and their ability to engage in mathematical practices, such as *communication*, or *reasoning and proof* (Common Core 2016, Kilpatrick *et al.* 2001). Specifically, these national standards imply that students should be active participants in constructing mathematical knowledge, through *discourse practices*. These practices engage students in elaborate, coherent, and focused classroom discourse, where the teacher and other students build on each other's responses, prompting each other to clarify and specify their ideas (Sadler 2006, Grossman *et al.* 2009b, Gomez Zaccarelli *et al.* 2018). They are intentionally designed to elicit students' thinking and reasoning to promote their mathematical conceptual knowledge (Schwartz *et al.* 2009).

To date, little research has been conducted about the effectiveness of established PD programs (Koellner *et al.* 2011, Borko *et al.* 2014). This is particularly true with regard to hybrid practice-based PD programs that use classroom videos as artefacts of practice (Barnett 2002, Dede *et al.* 2009, Santagata and Yeh 2014). Moreover, there is lack of research that examines classroom implementation of practices taught in a PD program, such as classroom discourse (Gomez Zaccarelli *et al.* 2018).

The current study aims to investigate the nature of changes in the way mathematics teachers who participated in a hybrid practice-based PD program practice classroom discourse in a real classroom setting. We examine these changes via the online coaching component of the program that occurred during one academic year, and in light of both teachers' reflections and coach's feedback. In what follows, we first discuss the scholarship on practice-based PD, specifically regarding the use of classroom videos and coaching; we then consider the specific core practice of classroom discourse.

Theoretical background

Practice-based professional development

Practice-based PD entails activities that are grounded in *authentic* tasks of teaching, in order to provide teachers with the opportunity to make connections between what they learn in the PD and their actual classroom instructional practice (Ball and Cohen 1999, Borko *et al.* 2010). Formats for PD programs include three important models: face-to-face programs, in which the PD facilitator and the teachers engage through in-person collaborative professional learning communities; online formats, defined as 'teacher learning experiences delivered partially or completely over the Internet' (Fishman *et al.* 2013, p. 427), which enable teachers to collaborate virtually; and hybrid programs that attempt to combine the strengths of the two previous forms, on the premise that the interactivity of online coaching is often not available in pure face-to face-teacher PD programs (Kleiman 2004, Dede *et al.* 2009).

The current study focuses on a hybrid PD program that combines facilitated group experiences in a summer institute, together with one-on-one online coaching during the academic year. The program uses activities to situate the PD within the teachers' experiences, for example, face-to-face rehearsals during the summer institute that simulate classroom instruction and provide the teachers with opportunities to practice skills such as facilitating classroom discussions with their peers as the

learners (Grossman *et al.* 2009a). It includes artefacts of practice such as video records of the teachers' lessons, which allow authentic engagement and analysis of practice in an online context (Santagata *et al.* 2007, Kazemi *et al.* 2009).

Use of classroom videos as artefacts of practice

Videos are uniquely capable of recording classroom variety and complexity for subsequent analysis (Brophy 2004, Sherin 2007, Jacobs *et al.* 2009). Video clips can be watched repeatedly and reviewed for different purposes. For example, in mathematics PD, classroom videos have enabled teachers to examine closely and discuss mathematics ideas, processes of thinking and learning, and instructional practices (e.g. Jacobs *et al.* 2009, van Es and Sherin 2010, Borko *et al.* 2014). Through video analysis, teachers improve their ability to comprehend instruction as a complex activity, and to cope with its complexity and uncertainty (Forzani 2014, Borko *et al.* 2017, Karsenty and Arcavi 2017). The premise is that by improving their ability to observe and analyse teaching practices, teachers will be able to make concrete connections to their own teaching, and to apply these practices in the settings of their own classroom (Kersting *et al.* 2010, Borko *et al.* 2011).

Coaching for teachers' professional learning

An important form of support, particularly for beginning teachers, is coaching (Marable and Raimondi 2007). Coaching is defined as 'one to-one support of a novice or less experienced practitioner by a more experienced practitioner [a coach]' (Hobson *et al.* 2009, p. 207). Studies have shown that coaching positively impacts teachers' classroom instructional practices, by improving the implementation of new practices and skills (Knight 2008).

Approaches to coaching for improving instruction are complex and nuanced (Poglinco and Bach 2004). They include observing and providing feedback through discussions before and after observations of teaching (Neufeld and Roper 2003, Gibbons and Cobb 2017, Parr and Hawe 2017), and providing concrete feedback while observing videos of the teacher's lessons (Santagata *et al.* 2007). In the PD program that is the focus of the current study, the coach provided feedback on videos that the teachers uploaded to an online platform. The program's approach to feedback incorporated praise for positive actions, explanations for why these actions are effective, questions about unclear actions, and suggestions about what could be improved for future instruction (Neubert and McNelis 1990). This feedback was designed to prepare the teacher for appropriate reactions to potential classroom events and support the teacher successfully to incorporate new practices independently (Knight 2008).

Focus on core-practice – facilitating classroom discourse

One of the core teaching practices emphasised in research and policy documents (e.g. Common Core, 2016) is facilitating classroom discourse. Research about classroom discourse reveals two main patterns of talk. The first refers to a recitation format (Nystrand *et al.* 1997), in which teachers control the discourse by asking closed-ended questions leading to pre-specified responses or 'right answers,' mostly via an I-R-E pattern of discourse (Mehan 1979). The I-R-E pattern involves *Initiation*, where the teacher poses questions, listens to students' answers (*Response*), and then assesses response correctness (*Evaluation*). This pattern does not enable opportunities for students to talk through their ideas, and therefore does not allow the teacher to gauge students' understanding or misunderstanding (Lemke 1990).

In contrast, more productive classroom discourse is based on sustained teacher-student dialogue, over a sequenced exchange of questions and answers that elicit students' thinking and reasoning (Erdogan and Campbell 2008, Tytler and Aranda 2015). Specifically, high-leverage classroom discussion in math lessons involves 'conceptual press', in which students are encouraged to give reasoning to mathematical concepts, in contrast to 'procedural press', which entails pushing for student responses

about procedural aspects, e.g. asking the students how to calculate an equation (Ball and Forzani 2009, Kazemi and Stipek 2009, Mevarech and Kramarski 2014). Cohen (2015) found that the math lessons of teachers who were rated as high in orchestrating discourse in their classroom often involved longer discussions with more opportunities for students to express and justify their solutions and their reasoning, mostly via open-ended questions initiated by the teachers.

Productive classroom discourse involves not just opportunities for students to talk, but also a process of 'uptake' of their responses, where teachers extend student ideas by incorporating those responses into subsequent questions (Collins 1982, Smith and Stein 2011, Smith and Hackling 2016). According to Hufferd-Ackles *et al.* (2004), productive classroom discourse in mathematics leads to a shift of questioning from the teacher as the only questioner to student-to-student initiated talk, where students are active participants in constructing their mathematical knowledge. This shift takes place as part of a process whereby the teachers encourage students to articulate their mathematical reasoning and propose mathematical ideas of their own. However, this does not imply that the teacher fully cedes responsibility for learning to the students. Rather, both student engagement and teacher intervention are required. The extent to which a dialogue is considered to be effective is dependent on the quality of interaction, i.e. the extent to which teachers expect students to think and formulate their own ideas, rather than simply repeat someone else's ideas (Alexander 2006).

There are particular moves that effective teachers use to lead productive classroom discussions, which can be initiated by either the teacher or the students themselves. These include moves for (a) initiating discussion, e.g. asking an open question; (b) building a discussion, e.g. connecting students' ideas to other students and building on them; and (c) closing the discussion to move on, e.g. students agree or disagree with other students using evidence (Smith and Stein 2011, Smith and Hackling 2016). Although the ultimate goal of facilitating classroom discourse is fostering students' productive participation, teachers are the ones who must create the necessary conditions for such participation. Therefore, it is important to provide teachers with PD experiences that will enable them to establish productive discourse in their classes (Crockett *et al.* 2009, Herbel-Eisenmann and Cirillo 2009).

The current study investigates the online coaching component of a hybrid practice-based PD program for mathematics teachers that is focused on supporting teachers' facilitation of classroom discourse through observing and reflecting on their video-recorded instructional practices. We address the following research questions:

(a) What differences in *discourse* instructional practices, if any, are associated with teachers' participation in online coaching?
(b) What is the relationship between both teachers' reflections and coach's feedback during the final coaching conversation, and changes in teachers' discourse instructional practices during online coaching?
(c) What is the relationship between both teachers' reflections and coach's feedback during the final coaching conversation, and level of teachers' discourse practices at the end of the online coaching year?

Method

The current study focuses on the *first year* of the *Hollyhock Fellowship Program*[1] (hereinafter, *HFP*) – a two-year professional development (PD) program that includes two 2-week summer institutes held on Stanford University campus, and online coaching sessions with experienced and knowledgeable practitioners throughout each academic year. More specifically, we address the online coaching that was provided in the first academic year, after the initial summer institute.

Participants

Research participants were 18 high-school mathematics teachers (3 males, 15 females) from throughout the United States who had been selected to participate as teaching fellows in one of the cohorts of the *HFP*. Each teaching fellow taught either Algebra or Geometry at a public, charter, or alternative high school. The program provided the teachers a travel stipend and a participation stipend for each academic year, as well as continuing education credits.

The fellows who were selected for the program were required to have between two and 7 years of teaching experience in one of the four foundational content areas of English, history/social science, maths, or science, and a teaching position at a low socio-economic status high school where more than 50% of students qualify for a free or reduced-price lunch. A small content-specific group of PD facilitators reviewed and ranked the fellowship applications in their content area. Then, all facilitators reviewed the overall school team rankings in an effort to create a diverse fellowship class in terms of school context, geography, and teacher experience.

The study was conducted after receiving approval from Stanford University's Institutional Review Board (#IRB-30,210). Four different Consent Forms were obtained from various participants in the research project prior to the *HFP*: Teacher Consent Form, Coach Consent Form, Parent Consent Form and Student Assent Form. The last two forms were translated into seven different languages and were obtained from the students who were captured in the videos and their parents.

The coach

One coach worked with the 18 teaching fellows. She has a Master's degree in mathematics and over 20 years of experience as a high-school mathematics teacher and site administrator in the US. She has served as a coach in the *HFP* since it began and is a member of the team that conducts the summer institutes. She also mentors early-career math teachers as they obtain their Professional Clear Credentials in Education or participate in other fellowship programs.

The hollyhock fellowship program

The *HFP* is a hybrid practice-based PD program. The current study focuses on the online coaching component of the program that occurred in the academic year that followed the first summer institute. In what follows, we briefly describe the first summer institute. We then describe the online coaching that followed.

Summer institute

The major portion of the first summer institute was the *Core Content Sessions* that focused on *facilitating productive discourse* in the math classroom. The sessions included a variety of activities such as partner talks, interactive lessons, and video-based discussions, to address the questions: what constitutes a 'discussion,' how discussion differs from other kinds of classroom talk, and why discussion is an essential activity structure and skill. During small group video-based content-area sessions, the teaching fellows learned and practice discussion skills. For example, they learned how and when to use discussion moves such as articulating, clarifying, critiquing, and building on each other's ideas; and how to identify moments in discussions when additional discussion moves could be applied to leverage student elaboration and argumentation.

The core content sessions also addressed how to teach a task-based unit, with emphasis on applying key discussion-based principles and practices to the unit plan. In addition, the teachers practiced leading discussions in rehearsal groups, where one or two of the teaching fellows serve as teachers, two or three other teaching fellows serve as observers, and the remaining teaching fellows act as students.

Online coaching

A key feature of the *HFP* was the coaching and support each teaching fellow received during the academic year following the summer institute. The coaching was comprised of online conversations that took place via audio or video tools, or both, during which the coach and teacher discussed video-recorded lessons from the teachers' classrooms that were uploaded to the program portal. During the academic year, teachers were required to upload *four* unedited videos of teaching on which they wished to receive feedback. The premise was that each video would include some classroom discourse. The videos were to be (preferably) at least 15 min in length. Teachers were encouraged to capture as many of the students in the video as possible, as long as they received signed consent from the students and their parents or guardians, allowing them to be included in the video. The teachers were responsible for ensuring video quality (audio and visual).

The online coaching focused on the facilitation of productive discourse in the context of a math classroom. The coaching sessions were designed to provide strategic and differentiated support for each teacher, by focusing on praise for the teacher's positive actions, questions about unclear actions, and suggestions for improvement. Prior to each session, the coach viewed the shared video and tagged specific moments for later discussion. Then, during the coaching conversations, the teaching fellow and the coach reflected on the video in real-time and discussed the tagged moments. In addition, discussion included planning goals for the next video to be recorded and uploaded for coaching. In their discussion of the fourth video, the teaching fellow and the coach also reflected on the past year in terms of teacher growth, celebrating 'successes' and considering 'needs' for the second summer institute. The goal of the coaching was that through the video uploads, the teaching fellows would demonstrate growth in the facilitation of class discussion, based on their PD experiences in the summer institute and the coaching conversations during the academic year.

The current study focuses on changes in teachers' instructional practices reflected in the videos that they uploaded. The online coaching involved the pathway described in Figure 1.

To examine changes in the teachers' instructional practices, we conducted a quantitative analysis of the first uploaded video that was captured after the initial summer PD, at the beginning of the academic year (September/October) and the fourth uploaded video that was captured towards the end of the academic year (March–May). In addition, we conducted a qualitative analysis of the fourth (final) conversation between the coach and each of the teaching fellows. These analyses are described below (*Method of analysis*).

Materials

To assess teachers' instructional quality, we used the 'Protocol for Language Arts Teaching Observation' (PLATO) observation tool. The protocol was designed to assess instructional quality in 15 minof time-sampling. PLATO has been used primarily for assessing the quality of language arts teaching (e.g. Grossman *et al.* 2013), but has also been modified to reliably score math instruction (Cohen 2015). PLATO includes six instructional elements that were found to be associated with student outcomes (for full description, see for example Grossman *et al.* 2011). Scores are given on a 4-point scale: 1 = Provides almost no evidence; 2 = Provides limited evidence; 3 = Provides evidence with some weaknesses; 4 = Provides consistent strong evidence. For this
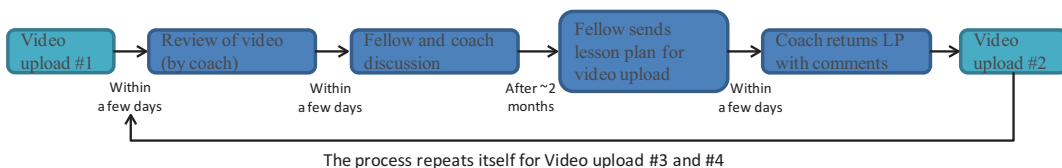

Figure 1. Timeline for online coaching.

study, we focus on the specific element of *classroom discourse* (hereafter, *CD*) that addresses the quality of instruction that facilitates productive discourse, through *opportunities* for student talk and *uptake* of student ideas within those opportunities. At the low-end (which represents a score of 1 or 2), teachers or students do not have opportunities to engage in mathematics-related discourse, or the discourse is closely teacher-directed with brief responses; at the high-end (which represents a score of 3 or 4), students have extended opportunities to engage in mathematics-related discussion that includes explanations, clarifications, or other high-level uptake of student ideas.

Reliability

To ensure reliability of video analysis, the first author, who conducted the ratings, completed training and certification as a PLATO rater. The training involved online modules that contained four short video clips that demonstrated the four score levels of the PLATO scale, as well as cognitive prompts that were designed to facilitate coding. For example, a cognitive prompt to support evidence towards a score of 1 for the CD element was '*Are responses to student ideas automatic or perfunctory (e.g. "good," "yes," "okay"), or does the teacher simply repeat what a student has said?*', while a cognitive prompt to support evidence towards a score of 4 for the CD element was '*Are there consistent/frequent moments of uptake and/or requests for clarification/elaboration on the part of the teacher or students?*'. Each module concluded with a check test that included a 15-min video clip for the rater to code. At the end, a reliability test was administrated. That test included five 15-min video clips, for which the rater scored each of the PLATO elements and provided brief notes to support the score. For each PLATO element, the rater's scores were compared with standardised scores, as determined by the PLATO developers. For elements that did not reach at least an 80% level of agreement, the rater reviewed the training modules and performed another reliability test. Certification was complete when the rater achieved 80% agreement for all PLATO elements.

Method of analysis

Quantitative analysis

As our goal in the current study was to assess changes in quality of CD from the first to fourth uploaded videos, we first searched for segments in which CD occurred. Although teachers were not explicitly requested to demonstrate discourse in their recorded video lessons, the premise was to present a discussion-based class. Indeed, all the videos involved at least 8 min of discussion between teacher and students, and/or between the students themselves. Because the PLATO instrument stipulates rating a 15-min segment, we used the following decision rule: first, we selected a segment in which we recognised a discussion; second, if the discussion was less than 15 min, then the segment ended at the conclusion of the discussion; finally, if the discussion was longer than 15 min, then the last 15 min were selected for analysis. After coding the 15-min segments of all videos, Paired-samples t-tests were conducted to compare the differences in PLATO-based ratings over time, and Chi-square tests of independence were performed to determine whether the likelihood of a change in teachers' PLATO-based ratings differed for teachers who received low ratings at the beginning of the year and those who received high ratings.

Qualitative analysis

We used an inductive approach to qualitative data analysis (Thomas 2006, Zhang and Wildemuth 2009), in the analysis of the fourth (final) conversation between the coach and each teacher. After transcribing the conversations, we read them carefully several times and identified recurring categories of issues related to CD. We distinguished between issues discussed by the teachers and issues discussed by the coach (which were labelled as types of feedback). These categories are detailed in the Findings section.

Because we were interested in investigating the relationship between the final coaching conversations and both the improvement in quality from the first to fourth videos and the quality of

discourse in the teacher's fourth video, we computed two indices of discourse practices for each teacher. The first index was based on change in the rating of CD, from the first to the last video: Teachers were separated into those who *improved* by 2 scores, *improved* by 1 score, and *remained* at the same level. (None of the teachers improved by 3 scores, from level 1 to level 4 of the PLATO observation tool.) The second index was based on the level of performance of CD in the last video (fourth lesson): a *high* level was assigned to teachers who demonstrated high-end performance of CD and a *low* level was assigned to teachers who demonstrated low-end performance of CD.

Results

Patterns of change in teachers' classroom discourse practices

To address the first research question, we provide a view of the change in instructional practices regarding the implementation of CD from the beginning to the end of the academic year. A paired-sample T-test with beginning and ending time as an independent variable was conducted, revealing a significant effect for time. This means that the quality of teachers' classroom discourse, as measured by PLATO, improved by the end of the academic year. Paired-samples t-tests were also conducted to compare the differences in scores over time for the two sub-elements of classroom discourse: uptake of student responses and opportunities for student talk. Findings revealed a significant improvement over time for both sub-elements (see Table 1).

Going beyond the average change over time, we observed the change in implementation of CD, separately for each of the 18 teachers who participated in the study. We conducted this analysis only for the CD element as a whole, not for the sub-elements.

As shown in Figure 2, half of the teachers (n = 9) increased the quality of CD in their classes, either a slight change representing continual low or high-end performance, i.e. improvement from rating1 to 2 or from rating 3 to 4; or more substantial change from low-end to high-end performance, i.e. from rating 2 to rating 3 or 4. For the remaining half of the teachers – those who demonstrated no change in their CD practice – 7 provided consistent strong evidence or evidence with some weaknesses in practice (i.e. ratings of 3 or 4 at both time points) and 2 demonstrated limited evidence of this practice (i.e. ratings of 2 at both time points).

To further explore the teachers' CD practice over time, we categorised the ratings into two levels: high (ratings of 3 or 4) and low (ratings of 1 or 2) and employed a Chi-square test of independence to determine whether the likelihood of a change in teachers' level of performance differed for teachers who received low ratings at the beginning of the year and those who received high ratings. Findings revealed a significant correlation, $\chi^2(1) = 6.92$, p < .01. See Table 2 for the *Chi-square test* results.

Table 1. Classroom discourse – uptake of student responses and opportunities for student talk: Means, standard deviations, t values, and Cohen's d scores by time (beginning and end of an academic year).

	Beginning (N= 18)		End (N= 18)			
	M	*SD*	*M*	*SD*	*d*	*t* (17)
Uptake of student responses	2.50	.99	3.17	.86	0.72	−3.69**
Opportunities for student talk	2.61	.98	3.33	.91	0.76	−3.20**
Total	2.50	.99	3.17	.86	0.72	−3.69**

(a) Scores ranged from 1 to 4, with higher scores indicating stronger evidence of this element in class.
(b) Cohen's *d* was calculated as following: Cohen's $d = (M_2 − M_1)/s_{pooled}$, where $s_{pooled} = \sqrt{[(s_1^2 + s_2^2)/2]}$ (M_1 represents the average beginning score; M_2 represents the average end score).
(c) T-values marked with asterisks (**) were significant at p < .01.
(d) According to the 'Protocol for Language Arts Teaching Observation' (PLATO) decision rules, when an element is broken down into two components, the first component is considered the most important, meaning that the first indicator (in the case of this study, the 'uptake of student responses') 'weighs more' when making a decision. If the two scores are adjacent (e.g. the first score is a 2 and the second score is a 3), then the score of the first indicator is chosen. If the two scores are not adjacent (e.g. a 2 and a 4), an average of the components is computed. In the current study, there were only adjacent scores or identical scores on the two components; therefore the statistics calculations for 'uptake' and 'total' are identical.
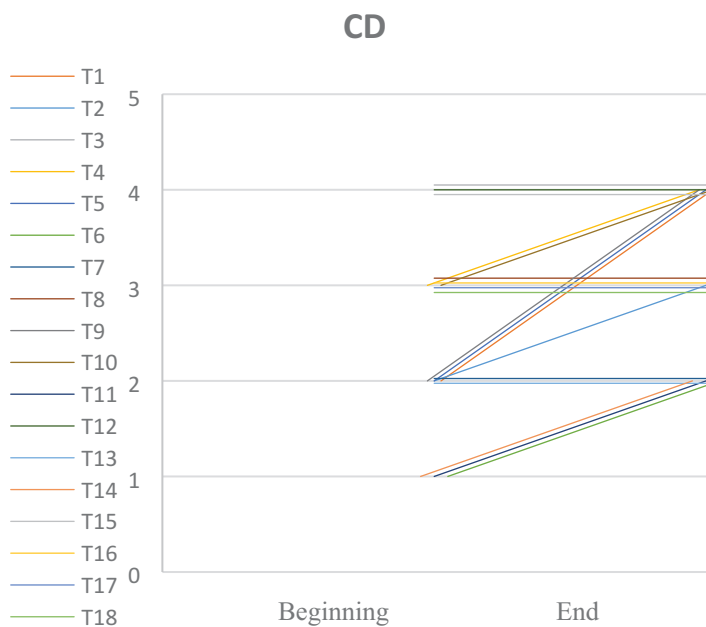
Figure 2. Change over time, in *CD*, by teacher, from the beginning to the end of an academic year.

Table 2. Results of Chi-square test and descriptive statistics for classroom discourse (CD), by level of performance at the beginning and at the end of an academic year.

Beginning	End	Classroom discourse		
		Low-rated [N (%)]	High-rated [N (%)]	Total [N (%)]
Low-rated		5 (55.6%)	4 (44.4%)	9 (100%)
High-rated		0 (0%)	9 (100%)	9 (100%)
Total		5 (27.8%)	13 (72.2%)	18 (100%)

Percentages were calculated for each row out of total teachers in each row.

As Table 2 shows, nearly half of the teachers who were low-rated at the beginning of the academic year were likely to show a high level of performance at the end of the academic year, while all of the teachers who were high-rated, were still likely to demonstrate this level of performance at the end of the academic year. In total, at the end of the academic year, about two thirds of the teachers were recognised as having a high level of *CD*. The comparison between low- and high-rated teachers, as well as the comparison between teachers who improved and teachers who demonstrated a similar level of CD is further examined in the following section.

The relationship between the coach's and teachers' reflections and teachers' classroom discourse practices

To address the second and third research questions, we investigated qualitatively the fourth (final) coaching conversation during the online coaching, in which the coach and the teachers reflected on teachers' instructional practices over the year. One teacher (T#6) left the program towards the end of the year, after uploading her last video-recorded lesson, and did not take part in the final

discussion with the coach. Thus, our analysis focuses on the 17 teachers who reflected on their experiences with the coach.

Examining patterns in the data, we noticed that 14 of the 17 teachers had a similar pattern of conversation. This pattern began with self-praise, in which the teachers were asked to reflect on the coach's initial open-ended question: 'Thinking of everything you have done this past year . . . what are you most pleased with . . . what do you feel is your greatest success?' Their reflection was then followed by the coach's feedback. The other three teachers (T#3, T#5, and T#9) did not reflect on their experiences spanning the entire year, rather only on the last video, expressing praise about a specific aspect of their teaching in that lesson. The remainder of the conversations focused on expectations for the second summer institute, and on general issues concerning teaching, such as final exams, the difficulty of covering the entire syllabus adequately, and the complexities of adolescent students. These general issues are beyond the scope of this paper and were not analysed.

We first consider the teachers' reflection in response to the coach's initial question (excluding teachers T#3, T#5, and T#9); then we examine the coach's feedback. For both teachers' reflections and coach's feedback, we made two comparisons: teachers who improved in their CD practice during the academic year were compared to those who remained at the same level of performance (to address the second research question) and teachers who were scored high on their last video were compared to those who were scored low on the same video (to address the third research question).

Teachers' reflections

Delving into teachers' reflections in response to the coach's initial question, we found that all but one of the teachers specifically discussed their CD implementation. We were able to discern broad consistencies across the conversations by identifying four categories of issues related to implementing CD: (1) explicit references to CD, by using the terms used in the *HFP*; (2) positive change towards implementing CD; and (3) impact of the *HFP* on their CD implementation. The fourth category of reflections addressed the teachers' need to continue improving their implementation of CD in practice.

Explicit reference refers to mentioning statements that were often heard with regard to facilitating CD during the *HFP*, such as moves to guide discussion and ensure accountability, e.g. asking students to revoice their ideas to the entire class, or allowing students to have their own discussions about mathematics without guiding the discussion or questions.

Table 3 presents the distribution of the categories of issues that appeared in teachers' reflections, organised by 1. teachers' level of performance in the last video (fourth lesson): high vs. low implementation of CD; and 2. level of change in teachers' implementation of CD, from the beginning (first video) to the end of the academic year (fourth [last] video). As Table 3 shows, there were a total of 39 comments made by the teachers in response to the coach's initial question. Seventeen of those comments (44%) were made by teachers who improved between the first and last lessons, and 22 (56%) were made by teachers who remained at the same level of performance. Thus, there was no significant difference in the number of comments made by teachers who improved and teachers who remained at the same level. When comparing the frequencies for each category, we detected this similarity more clearly. For example, 18% of the comments (7 out of a total of 39) were made by teachers who improved and explicitly referenced CD, whereas 21% of the comments (8 out of a total of 39) were made by teachers who remained at the same level and explicitly referenced CD. In other words, there was no difference in the frequency with which the issues related to CD were discussed by the teachers who demonstrated improvement in implementation of discourse in their classes from the first to the fourth lesson, compared to those whose practice remained at the same level.

With regard to the relationship between the types of issues discussed by the teachers and their level of performance at the end of the academic year, we found differences in both frequency and specificity between teachers who were rated high vs. low. As one example, all of the teachers (100%) who were rated high on CD implementation on their last video explicitly mentioned the terms that were used during the PD program when asked to reflect on their greatest success, either referring to a specific

Table 3. Distribution of categories mentioned by the teachers during reflective conversations with the coach, by level of performance of CD in fourth lesson and level of change from first to fourth lesson.

			Level of change from first to fourth lesson			
		Category	Improved by 2 scores (N = 3)	Improved by 1 score (N = 6)	Remained the same level (N = 9)	Total (%)
Level of performance of CD in fourth lesson	High level (N = 13)	Explicit reference to CD (using the terms of the PD)	3	3	7	13 (100%)
		Positive change in CD implementation	1	3	6	10 (77%)
		Impact of the PD	0	1	2	3 (23%)
		Need to continue improving CD implementation	2	2	1	5 (38%)
	Low level (N = 4)	Explicit reference to CD (using the terms of the PD)	N/A	1	1	2 (50%)
		Positive change in CD implementation	N/A	1	2	3 (75%)
		Impact of the PD	N/A	0	1	1 (25%)
		Need to continue improving CD implementation	N/A	0	2	2 (50%)
Total			6 (16%)	11 (28%)	22 (56%)	39 (100%)

While the column percentages represent the total percentages of the comments made by teachers who improved or remained at the same level, percentages for each row were calculated from the total of high-rated teachers (upper part of the table) or low-rated teachers (lower part of the table) who completed the program. As each teachers' comment was assigned to only one category, the frequencies in each row represent both the number of teachers and the number of comments made by them.

move or to ensuring accountability. For example, T#8 began her reflections by describing the accountability in her class[2]:

> I think the biggest thing for me is recognizing that students learn best **when they are in charge of their own learning**, so I'm not explaining something to them … they are really creating their understanding of the concepts …

Then, she deepened her explanation by explicitly specifying a CD move as she continued:

> … students were constructing their understanding **by taking control of the conversations**, so if a question comes up, it is another student that is answering the question … students are asked to really be in charge in the classroom, rather than me … I feel that I talk a lot less on a typical day, which is nice. Because it's not about me. It is about them doing the work …

By contrast, only half of the low-rated teachers explicitly referred to implementing CD in their last video. For example, T#11 reflected on utilising a task-based unit, which was a tool addressed in the PD summer institute as a way to implement productive classroom discussions. However, he referred to this practice as related to the curriculum and did not refer at all to discourse.

Further, we found that 77% of the high-rated teachers referred to a positive change, similarly to 75% of the low-rated teachers. Yet, there was a difference in the specificity of teachers' statements. For example, T#13, who was low-rated, said without adding any detail:

> … It's still a lot of work in progress, but I'm getting more discussion …

In contrast, T#12, who was high-rated, referred to a positive change in a feature of CD, with regard to providing opportunities for students to compare and contrast ideas:

> What I found good was that towards the end of the year, my students became more comfortable … **in terms of talking about it and making comparisons. I could step back to let the students figure it out …** I let the students talk it through, do it together as a team, and then in the whole group we came back together and then discussed what they had noticed, or what they wondered …

Although the teachers were asked to reflect on their greatest success, 38% of the high-rated teachers and 50% of the low-rated teachers referred to their need to continue improving their CD implementation. In addition to difference in frequency, there was a difference in the specificity of the teachers' statements. For example, T#7, who was low-rated, referred to the need to continue improving as follows:

> I don't know … I thought that I improved in some areas … in terms of discussions, it is not something I have ever done before. I think there is still a lot of room for improvement … But it's better. I have never really implemented discussion in my classroom …

T#1, who was high-rated, also referred to her need to improve; however, she was more specific about her need to improve her ability to foster discussions that were student-led rather than teacher-directed. She attributed this insight to what she had learned from the coach.

> I think the biggest success that I feel is being super-intentional and feeling the accountability … I do feel that there is still a lot of room for improvement with the discussion piece … I think I would like … you have talked about this before … **to have the discussion more student-led** … I kind of thought that they were going to get more into it … I anticipated that someone would at least say: no, I think it's something else, so I sort of redirected the discussion to the way that I wanted it to go …

Finally, we found a similar percentage of high-rated and low-rated teachers (about 25%) who referred to the impact of PD on their CD implementation. Yet, there were differences in the specificity of the teachers' statements. T#11, who was low-rated, expressed the contribution of the *HFP* summer institute as follows:

> This [facilitating discussions in class] is definitely new. Some of this was [discourse] language dependent … in my core class we sort of present ideas and there is a conversation about agreement or disagreement, but this year it is a little bit more rule based, like what kind of conversation we can have

Whereas T#11 referred briefly to the change in formality of discussion, T#18, who was high-rated, referred to a much broader change in the implementation of CD, as a result of the *HFP*:

> The moves [of discussion] we've learned … I understand what is behind them. When we began this, I did not understand the power of discussion. When we saw the video of the students performing this discussion with each other [in the PD summer institute] … I began to see the difference. Once I did most of the work. Today students do much of the work … they respond by "**I agree because**", "**I disagree because** ", "**I noticed that … ".** I don't have to say the moves; they do it by their own.

In summary, we discovered a relationship between teachers' change with respect to classroom discourse practices, and the types of issues they discussed during the last coaching conversation, reflecting their greatest successes. The relationship was manifested by a greater *frequency* of explicit references to terms that were used during the PD summer institute and the online coaching by the high-rated teachers when compared to the low-rated teachers, and a greater *frequency* of terms related to the need to continue improving CD implementation by the low-rated teachers. There also was greater *specificity* in the high-rated teachers' comments about all four categories of issues.

Coach's feedback

The coaching component of the PD used a praise and polish format for feedback, both when reflecting on each teacher's instructional practices, and when responding to teachers' reflections. We noticed, however, that in the fourth (and last) coaching conversation for this year, which was the conversation analysed in the current study, most of the coach's feedback (59%) referred only to the praise. We classified these comments as Positive. Six per cent of the coach's feedback referred only to ways of polishing or improving the teacher's practice, which we classified as Critical

feedback. Further, 35% of the coach's feedback referred to a mixture of praise and polish, which we addressed as 'Positive + Critical' feedback. Following is an example of feedback that started out positive and became critical:

> You are doing a **great job** [**Positive**: praise] by asking the students for evidence, pushing on their thinking and having them explain themselves … one of the things **you need to work on** [**Critical**: polish] is taking students' ideas and launching them into a deeper discussion … **your next step is** [**Critical**: polish] to think about what question I need to ask in order to move the discussion further …

Table 4 presents the distribution of each type of feedback organised by 1. teachers' level of performance in the last video (fourth lesson): high vs. low implementation of CD; and 2. level of change in teachers' implementation of CD, from the beginning (first video) to the end (fourth [last] video) of the academic year.

Reviewing Table 4, we found a difference in the nature of feedback given to teachers who improved, when compared to the feedback given to teachers who remained at the same level of CD performance. Specifically, 5 of the 6 feedback statements that included *a mixture of positive and critical* comments were given to teachers who demonstrated no change in their CD implementation over the academic year; only one teacher who demonstrated an increase in CD implementation received this kind of feedback. In contrast, 7 of the 10 feedback statements that were *positive*, were given to teachers who demonstrated an increase in their CD implementation, compared to only 3 that were given to teachers who demonstrated no change. (As there was only one comment that was exclusively critical, we do not consider that comment further in this section.) Regarding the second comparison between teachers who received high scores on their last videos and those who scored low, teachers who were scored *high* received mostly positive feedback (8 of 13), while teachers who were scored *low* received the same number of positive comments and positive + critical comments (2 of 4).

The following two examples illustrate these two comparisons; the first refers to T#9, who demonstrated a significant improvement, as her rating improved from a low score (2) to the highest score (4). She was given positive feedback by the coach, with regard to her last video, as well as her progress throughout the year:

> I was thrilled when I saw your video. You stepped out of the box a little bit and tried something new … You did a great job at the beginning about setting your norms, setting your expectations, modeling what you wanted to hear … **You have tried something new each time**, and the last video was a huge step for you, because you really let the students discuss the mathematics and come up with the content that they are going to talk about and you facilitated that. That was awesome. I was so proud of you. That was really, really good …

Table 4. Distribution of types of feedback given by the coach during conversations with the teachers, by teachers' level of performance of CD in fourth lesson and level of change from first to fourth lesson.

			Level of change from first to fourth lesson			
		Type of feedback	*Improved* by 2 scores (N = 3)	*Improved* by 1 score (N = 6)	*Remained* at same level (N = 9)	Total (%)
Level of performance of CD in fourth lesson	High level (N = 13)	Positive	3	2	3	8 (62%)
		Critical	0	0	1	1 (8%)
		Positive+ Critical	0	1	3	4 (31%)
	Low level (N = 4)	Positive	N/A	2	0	2 (50%)
		Critical	N/A	0	0	0 (0%)
		Positive + Critical	N/A	0	2	2 (50%)
	Total (N = 17)	Positive	3	4	3	10 (59%)
		Critical	0	0	1	1 (6%)
		Positive + Critical	0	1	5	6 (35%)

Percentages were calculated out of total teachers who completed the program, rated as high (top part of the table), low (middle part of the table), or total (bottom part of the table). Because each teacher received feedback that could be assigned to only one feedback type, the frequencies in the table represent both the number of teachers and the number of feedback comments received by the coach.

The second example refers to T#13, who did not improve from her first to her fourth lesson and was low-rated (2) in both lessons. She received a mixture of positive and critical feedback:

> Going to the video, your students are very comfortable with presenting [ideas in the conversation] … My push to you is … procedure is so important in mathematics, because if they do not have the procedure they cannot solve the problem … but … try to bring in questions that have the students think about the concept, the application, the big questions. For example, if you are solving quadratic equations, what does that actually mean … start making connections to other aspects of the unit.

In sum, the qualitative analysis regarding the nature of changes in teachers' classroom discourse instructional practices examined the type of issues discussed in teachers' reflections and the coach's feedback during the final coaching conversation. In examining these issues, we considered two comparisons: (1) teachers who improved in their practice of CD during the academic year were compared to those who remained at the same level of performance; and (2) teachers who were rated high on their last video were compared to those who were rated low on the last video. Teachers' reflections revealed more differences between high-rated and low-rated teachers than between teachers who improved and those whose ratings remained the same. The high-rated teachers' reflections included more explicit references to CD and fewer references to their need to continue improving CD, compared to the low-rated teachers. In addition, the statements of the high-rated teachers were more detailed and more specific, with respect to all issues discussed. We also found a relationship between the issues discussed in the coach's feedback and the change, or lack thereof, in teachers' instructional practices. Specifically, teachers who improved in their CD implementation over the year, as well as teachers who were high-rated on their last video, received fewer comments that included a mixture of positive and critical feedback than did the teachers who remained at the same level and those who were low-rated on their last video.

Discussion and conclusions

The current study investigates the effect of the *HFP* – a hybrid practice-based professional development (PD) program – on teaching practices focused on facilitation of productive discourse in mathematics classes. The study contributes to both research and practice in the field of professional development for teachers, particularly to the growing literature about the importance of high-quality practice-based PD programs (Desimone and Garet 2015, Gaudin & Chalies, 2015) and the need to investigate the effectiveness of such programs (Borko *et al.* 2010, Gibson and Ross 2016, Gomez Zaccarelli *et al.* 2018).

The study adds to the limited research literature on the effectiveness of hybrid PD programs (Barnett 2002, Dede *et al.* 2009). The findings that: (a) half of the teachers improved in their classroom instructional discourse practices over the course of the year of coaching sessions and no teachers' scores decreased; and (b) the teachers rated high on CD expressed more explicit, detailed and specific references to CD in their final coaching conversation offer some evidence for the value of providing teachers with sustained support throughout the year and, more specifically, the online coaching component of the hybrid *HFP*, for facilitating positive change in teachers' classroom discourse.

The study also adds to the literature on the effectiveness of classroom video as an artefact of practice for PD focused on improving teachers' practice (Borko *et al.* 2008, Santagata and Yeh 2014). Several of coach's comments during the final coaching conversation referred specifically to instructional strategies on the teacher's video. These comments suggest that she found the video to be a useful vehicle for providing both positive feedback and suggestions for improvement. The teachers also noted the value of video as a tool in professional development, as evidenced by T#18's comment about the video of a student discussion used in the summer institute. As these examples indicate, video records of classroom lessons provided an opportunity for the coach and teachers to analyse the CD practices introduced in the summer institute and reinforced in the coaching sessions.

Methodologically, the study contributes to previous studies regarding the effectiveness of classroom video as a research tool to study teacher change (Borko *et al.* 2008, Jacobs *et al.* 2009, Seidel *et al.* 2009) and the usefulness of the PLATO observation protocol to code these changes (Grossman *et al.* 2013). Strong mathematics CD involves teachers' moves for *uptake* of students' responses, as well as moves that extend *opportunities* for engaging students in mathematics-related discussions (Kazemi and Stipek 2009, Cohen 2015). Our study demonstrates that these CD moves can be observed and measured in mathematics classrooms using the PLATO tool. Moreover, the teachers' reflections and the coach's feedback, which corresponded with the positive change captured by PLATO, adds to the reliability of the protocol as a tool for assessing discourse in mathematics lessons.

The current study also makes several practical contributions. First, although the study involved a relatively small number of mathematics teachers, most of them demonstrated positive changes in their *observed* teaching practice, as well as positive and specific comments in *reflective conversations* regarding their practice. These changes illustrate the potential of practice-based PD for achieving the goal of connecting the *theory* of classroom discourse and its implementation in *practice* (Desimone and Garet 2015).

Second, early career teachers often feel unprepared to apply what they learned about teaching during their teacher preparation programs in actual practice in the school setting (Liston *et al.* 2006, Ingersoll and Perda 2009, Banks *et al.* 2014). The current study responds to the challenge of helping these teachers develop and strengthen the types of knowledge and skills that are necessary for effective teaching in practice, through professional support and coaching before and during an academic year (National Commission on Teaching and America's Future 2003, Borko *et al.* 2010, Gibson and Ross 2016).

Third, according to mathematics teaching standards, facilitating discourse in mathematics is fundamental to mathematics learning (e.g. Common Core., 2016); however, it is frequently absent from typical math classrooms (Sadler 2006). This study contributes to our understanding of ways to improve mathematics classroom discourse, as it demonstrates the effectiveness of a PD program focused on classroom discourse instructional practices for helping early career teachers learn to enact specific instructional moves that effective teachers use to facilitate productive mathematics classroom discussions (Hufferd-Ackles *et al.* 2004).

Research limitation and future study

In the present study, we examined changes in teachers' CD instructional practices over one academic year, which followed their participation in a summer institute that included sessions focused on facilitating productive discourse. These data did not allow us to separately investigate the effects of the two features of the hybrid *HFP* – the summer institute and the online coaching sessions. Additional types of data would shed light on these effects, for example, baseline videos collected and uploaded before the summer institute, and interviews with the teachers immediately after the summer institute and at the conclusion of the academic year.

For future study, we also suggest investigating a PD program that only includes the online feature of the current hybrid *HFP*, and its potential to address the challenge of scaling up PD programs (e.g. Koellner *et al.* 2011). Referring specifically to the online feature of the *HFP*, we noticed a relationship between the coach's feedback and teachers' reflection, mainly for the teachers who were highly rated in their last video-recorded lesson. Yet, we did not directly investigate the relationship between *specific* issues discussed in coaching conversations during the online coaching, and *specific* changes in individual teachers' discourse instructional practices. We suggest exploring this relationship in further study and including a comparison of the relationship between coaching conversations and instructional practices for the teachers whose discourse practices improved and the low-rated teachers who were unable to consistently implement the discourse practices addressed in the summer institute and online coaching.

Notes

1. https://cset.stanford.edu/pd/hollyhock.
2. Bolded text identifies terms explicitly used during the *HFP* summer institute and online coaching.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Zehavit Kohen   http://orcid.org/0000-0003-2084-4780

References

Alexander, R., 2006. *Towards dialogic teaching*, 68 (4), 411–425. York, UK: Dialogos.
Ball, D.L. and Cohen, D.K., 1999. Developing practice, developing practitioners: toward a practice-based theory of professional education. *In*: G. Sykes and L. Darling-Hammond, eds. *Teaching as the learning profession: handbook of policy and practice*. San Francisco: Jossey Bass, 3–32.
Ball, D.L. and Forzani, F., 2009. The work of teaching and the challenge for teacher education. *The journal of teacher education*, 60 (5), 497–511. doi:10.1177/0022487109348479.
Banks, T., Jackson, D., and Harper, B., 2014. Responding to the call to prepare highly effective teacher candidates in the united states: the curriculum redesign effort in advancing teacher education. *Higher education studies*, 4 (2), 9–18. doi:10.5539/hes.v4n2p9.
Barnett, M., 2002. *Issues and trends concerning electronic networking technologies for teacher professional development: A critical review of the literature*. Paper presented at the American Educational Research Association. New Orleans, LA.
Borko, H., *et al*., 2008. Video as a tool for fostering productive discussions in mathematics professional development. *Teaching and teacher education*, 24 (2), 417–436. doi:10.1016/j.tate.2006.11.012.
Borko, H., *et al*., 2011. Using video representations of teaching in practice-based professional development programs. *ZDM*, 43 (1), 175–187. doi:10.1007/s11858-010-0302-5.
Borko, H., *et al*., 2014. Facilitating video-based professional development: planning and orchestrating productive discussions. *In*: Y. Li, E. A. Silver, & S. Li, Eds.*Transforming mathematics instruction*. Cham: Springer, 259–281.
Borko, H., *et al*., 2017. The role of video-based discussion in model for preparing professional development leaders. *International journal of STEM education*, 4 (1), 29. doi:10.1186/s40594-017-0090-3.
Borko, H., Jacobs, J., and Koellner, K., 2010. Contemporary approaches to teacher professional development. *International encyclopedia of education*, 7 (2), 548–556.
Brophy, J., Ed., 2004. *Advances in research on teaching, Vol. 10: using video in teacher education*. Oxford, UK: Elsevier.
Cohen, J., 2015. Challenges in identifying high-leverage practices. *Teachers college record*, 117 (7), 1–41.
Collins, J., 1982. Discourse style, classroom interaction and differential treatment. *Journal of reading behavior*, 14 (4), 429–437. doi:10.1080/10862968209547468.
Common Core State Standards Initiative. 2016. *Common core state standards for mathematics*. Available from: http://www.corestandards.org/wp-content/uploads/Math_Standards1.pdf [Accessed 18 Dec 2019].
Crockett, M.D., *et al*., 2009. Exploring discourse-based assessment practice and its role in mathematics professional development. *Professional development in education*, 35 (4), 677–680. doi:10.1080/13674580802594175.
Dede, C., *et al*., 2009. A research agenda for online teacher professional development. *Journal of teacher education*, 1 (60), 8–19. doi:10.1177/0022487108327554.
Desimone, L.M. and Garet, M.S., 2015. Best practices in teachers' professional development in the United States. *Psychology, society, & education*, 7 (3), 252–263. doi:10.25115/psye.v7i3.
Erdogan, I. and Campbell, T., 2008. Teacher questioning and interaction patterns. *International journal of science education*, 30 (14), 1891–1914. doi:10.1080/09500690701587028.
Fishman, B., *et al*., 2013. Comparing the impact of online and face-to-face professional development in the context of curriculum implementation. *Journal of teacher education*, 64 (5), 426–438. doi:10.1177/0022487113494413.
Forzani, F.M., 2014. Understanding "Core Practices" and "Practice-Based" teacher education learning from the past. *Journal of teacher education*, 65 (4), 357–368. doi:10.1177/0022487114533800.
Gaudin, C. and Chaliès, S., 2015. Video viewing in teacher education and professional development: A literature review. *Educational research review*, 16, 41–67. doi:10.1016/j.edurev.2015.06.001

Gibbons, L.K. and Cobb, P., 2017. Focusing on teacher learning opportunities to identify potentially productive coaching activities. *Journal of teacher education*, 68, 411–425. doi:10.1177/0022487117702579

Gibson, S.A. and Ross, P., 2016. Teachers' professional noticing. *Theory into practice*, 55 (3), 180–188. doi:10.1080/00405841.2016.1173996.

Gomez Zaccarelli, F., *et al*., 2018. Learning from professional development: A case study of the challenges of enacting productive science discourse in the classroom. *Professional development in education*, 44 (5), 721–737. doi:10.1080/19415257.2017.1423368.

Grossman, P., *et al*., 2009a. Teaching practice: A cross-professional perspective. *Teachers college record*, 111 (9), 2055–2100.

Grossman, P., *et al*., 2009b. *Development of the protocol for language arts teaching observation (PLATO)*. In Annual meeting of the American Educational Research Association, San Diego, CA.

Grossman, P., *et al*., 2013. Measure for measure: the relationship between measures of instructional practice in middle school English language arts and teachers' value-added scores. *American journal of education*, 119 (3), 445–470. doi:10.1086/669901.

Grossman, P., Cohen, J., and Brown,, 2011. *Classroom practices associated with value-added measures of teacher effectiveness in english language arts*. Paper presented at the American Education Research Association's annual meeting in New Orleans. April 10, 2011.

Herbel-Eisenmann, B. and Cirillo, M., Eds., 2009. *Promoting purposeful discourse: teacher research in mathematics classrooms*. Reston, VA: NCTM.

Hobson, A.J., *et al*., 2009. Mentoring beginning teachers: what we know and what we don't. *Teaching and teacher education*, 25 (1), 207–216. doi:10.1016/j.tate.2008.09.001.

Hufferd-Ackles, K., Fuson, K.C., and Sherin, M.G., 2004. Describing levels and components of a math-talk learning community. *Journal for research in mathematics education*, 35 (2), 81–116. doi:10.2307/30034933.

Ingersoll, R.M. and Perda, D., 2009. *The mathematics and science teacher shortage: fact and myth (CPRE Rep. No. RR-62)*. Philadelphia: Consortium for Policy Research in Education.

Jacobs, J., Borko, H., and Koellner, K., 2009. The power of video as a tool for professional development and research: examples from the problem-solving cycle. *In*: T. Janik and T. Seidel, eds. *The power of video studies in investigating teaching and learning in the classroom*. Munster: Waxmann Publishing, 259–273.

Karsenty, R. and Arcavi, A., 2017. Mathematics, lenses and videotapes: a framework and a language for developing reflective practices of teaching. *Journal of mathematics teacher education*, 20, 433–455. doi:10.1007/s10857-017-9379-x

Kazemi, E. and Franke, M.L., 2004. Teacher learning in mathematics: using student work to promote collective inquiry. *Journal of mathematics teacher education*, 7 (3), 203–235. doi:10.1023/B:JMTE.0000033084.26326.19.

Kazemi, E., Franke, M.L., and Lampert, M., 2009. *Developing pedagogies in teacher education to support novice teachers' ability to enact ambitious instruction*. Paper presented at the annual meeting of the Mathematics Education Research Group of Australasia, Wellington, New Zealand.

Kazemi, E. and Stipek, D., 2009. Promoting conceptual thinking in four upper-elementary mathematics classrooms. *Journal of education*, 189 (1–2), 123–137. doi:10.1177/0022057409189001-209.

Kersting, N.B., *et al*., 2010. Teachers' analyses of classroom video predict student learning: further explorations of a novel measure of teacher knowledge. *Journal of teacher education*, 61, 172–181. doi:10.1177/0022487109347875

Kilpatrick, J., Swafford, J., and Findell, B., 2001. *Adding it up: helping children learn mathematics*. Washington, DC: National Academy Press.

Kleiman, G.M., 2004. Meeting the need for high quality teachers: E-learning solutions. Commissioned White Paper for the U.S. Department ofEducation Secretary's Leadership Summit on e-Learning and the U.S. Department of Education National Technology Plan.

Knight, J., 2008. *Coaching: approaches and perspectives*. Thousand Oaks, CA: Corwin Press.

Koellner, K., Jacobs, J., and Borko, H., 2011. Mathematics professional development: critical features for developing leadership skills and building teachers' capacity. *Mathematics teacher education and development*, 13 (1), 115–136.

Lemke, J., 1990. *Talking science: language, learning, and values*. Norwood, NJ: Ablex Publishing Coorporation.

Liston, D., Whitcomb, J., and Borko, H., 2006. Too little or to much: teacher preparation and the first years of teaching. *Journal of Teacher Education*, 57, 351–358. doi:10.1177/0022487106291976

Loucks-Horsley, S., *et al*., 2010. *Designing professional development for teachers of science and mathematics*. 3rd. Thousand Oaks, CA: Corwin.

Marable, M.A. and Raimondi, S.L., 2007. Teachers' perceptions of what was most (and least) supportive during their first year of teaching. *Mentoring & Tutoring*, 15 (1), 25–37. doi:10.1080/13611260601037355.

Mehan, H., 1979. *Learning lessons*. Cambridge, MA: Harvard University Press.

Mevarech, Z. and Kramarski, B., 2014. critical maths for innovative societies: the role of metacognitive pedagogies. OECD Publications. https://www.oecd.org/publications/critical-maths-forinnovative-societies-9789264223561-en.htm

National Commission on Teaching and America's Future, 2003. *No dream denied: A pledge to America's children*. Washington, DC: Author.

Neubert, G.A. and McNelis, S.J., 1990. Peer response: teaching specific revision suggestions. *The english journal*, 79 (5), 52–56. doi:10.2307/818375.

Neufeld, B. and Roper, D., 2003. *Coaching: A strategy for developing instructional capacity, promises and practicalities*. Washington, DC: Aspen Institute Program on Education and Annenberg Institute for School Reform.

Nystrand, M., *et al.*, 1997. *Opening dialogue: understanding the dynamics of language and learning in the English classroom*. New York, NY: Teachers College Press.

Parr, J.M. and Hawe, E., 2017. Facilitating real-time observation of, and peer discussion and feedback about, practice in writing classrooms. *Professional development in education*, 43 (5), 709–728. doi:10.1080/19415257.2016.1241818.

Poglinco, S.M. and Bach, A.J., 2004. The heart of the matter: coaching as a vehicle for professional development. *Phi Delta Kappan*, 85 (5), 398–400. doi:10.1177/003172170408500514.

President's Council of Advisors on Science and Technology (PCAST), 2010. Prepare and inspire: K–12 education in science, technology, engineering, and math (STEM) for America's future: executive report. Executive Office of the President, President's Council of Advisors on Science and Technology. Available from: https://nsf.gov/attachments/117803/public/2a–Prepare_and_Inspire–PCAST.pdf

Sadler, T.D., 2006. Promoting discourse and argumentation in science teacher education. *Journal of science teacher education*, 17 (4), 323–346. doi:10.1007/s10972-006-9025-4.

Santagata, R. and Yeh, C., 2014. Learning to teach mathematics and to analyze teaching effectiveness: evidence from a video-and practice-based approach. *Journal of mathematics teacher education*, 17 (6), 491–514. doi:10.1007/s10857-013-9263-2.

Santagata, R., Zannoni, C., and Stigler, J.W., 2007. The role of lesson analysis in preservice teacher education: an empirical investigation of teacher learning from a virtual video-based field experience. *Journal of mathematics teacher education*, 10, 123–140. doi:10.1007/s10857-007-9029-9

Schwartz, Y., *et al.*, 2009. Talking science. *The science teacher*, 76 (5), 44–47.

Seidel, T., *et al.*, 2009. LUV and Observe: two projects using video to diagnoses teachers' competence. *In*: T. Janik and T. Seidel, eds. *The power of video studies in investigating teaching and learning in the classroom*. Munster: Waxmann Publishing, 243–258.

Sherin, M.G., 2007. The development of teachers' professional vision in video clubs. *In*: R. Goldman, *et al.*, eds. *Video research in the learning sciences*. Mahwah, NJ: Lawrence Erlbaum Associates, 383–390.

Smith, M.S. and Stein, M.K., 2011. *Five practices for orchestrating productive mathematics discussions*. Reston, VA: National Council of Teachers of Mathematics.

Smith, P.M. and Hackling, M.W., 2016. Supporting teachers to develop substantive discourse in primary science classrooms. *Australian journal of teacher education*, 41 (4), 10. doi:10.14221/ajte.

Thomas, D.R., 2006. A general inductive approach for analyzing qualitative evaluation data. *American journal of evaluation*, 27 (2), 237–246. doi:10.1177/1098214005283748.

Tytler, R. and Aranda, G., 2015. Expert teachers' discursive moves in science classroom interactive talk. *International journal of science and mathematics education*, 13 (2), 425–446. doi:10.1007/s10763-015-9617-6.

van Es, E.A. and Sherin, M.G., 2010. The influence of video clubs on teachers' thinking and practice. *Journal of mathematics teacher education*, 13, 155–176. doi:10.1007/s10857-009-9130-3

Zhang, Y., & Wildemuth, B.M. (2009). Qualitative analysis of content. In B. Wildemuth (Ed.), Applications of Social Research Methods to Questions in Information and Library Science (pp. 308–319). Westport, CT: Libraries Unlimited.