2015

Professional Development at the Local Level: Necessary and Sufficient Conditions for Critical Colleagueship

Tara Kintz
Michigan State University

John Lane
Michigan State University

Amelia Gotwals
Michigan State University

Dante Cisterna
Pontifical Catholic University of Chile

Follow this and additional works at: <https://digitalcommons.unl.edu/natrespapers>

Part of the [Natural Resources and Conservation Commons](#), [Natural Resources Management and Policy Commons](#), and the [Other Environmental Sciences Commons](#)

Kintz, Tara; Lane, John; Gotwals, Amelia; and Cisterna, Dante, "Professional Development at the Local Level: Necessary and Sufficient Conditions for Critical Colleagueship" (2015). *Papers in Natural Resources*. 952.
<https://digitalcommons.unl.edu/natrespapers/952>

This Article is brought to you for free and open access by the Natural Resources, School of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Papers in Natural Resources by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Professional development at the local level: Necessary and sufficient conditions for critical colleagueship

Tara Kintz ^{a,*}, John Lane ^{a,1}, Amelia Gotwals ^{a,1}, Dante Cisterna ^{b,1}

^a Michigan State University, 620 Farm Lane, Erickson Hall, East Lansing, MI 48843, USA

^b Pontifical Catholic University of Chile, Facultad de Educacion, Av. Vicuna Mackenna 4860, Macul, Santiago, 7820436, Chile

HIGHLIGHTS

- Investigation of conditions associated with teachers' in-depth discussions.
- Focus on teacher communities of inquiry in a professional development initiative.
- Using unique methodological approach: Qualitative Comparative Analysis.
- Single purpose was a necessary condition associated with in-depth discussion.
- Coach questions and connecting theory and practice were also associated conditions.

ARTICLE INFO

Article history:

Received 3 December 2014

Received in revised form

7 May 2015

Accepted 17 June 2015

Available online 6 July 2015

Keywords:

Communities of practice
Professional development
Qualitative comparative analysis
Teacher inquiry

ABSTRACT

This paper examines factors that contributed to critical conversations in teacher communities of inquiry (CI) as part of a statewide professional development initiative in the United States. Based on a three-year mixed method design, we use qualitative comparative analysis to investigate the influence of combinations of conditions on the depth of discussion. Results suggest that there were three conditions associated with the extent to which CI members engaged in discussions with substantive interaction and reflection: a clear purpose, coach questioning, and the connection of theory to practice. The findings contribute to the understanding of effective reform implementation in different contexts.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

High quality teacher professional development (PD) has the potential to promote: increased student achievement (e.g., Desimone, Smith, Hayes, & Frisvold, 2005), high-quality schools (e.g., Borko & Putnam, 1995; Desimone, 2009; Talbert, McLaughlin, Rowan, 1993), and effective policy implementation (McIntyre & Kyle, 2006). Research conducted in the United States, however, shows that the quality and effectiveness of PD programs vary considerably (Yoon, Duncan, Lee, Scarloss, & Shapley, 2007). Many efforts to promote teacher learning have been inconsistent or ineffective (Coburn, 2001; McLaughlin, 1991) and reforms often fall

short of providing meaningful PD that promotes change (Darling-Hammond, Wei, Andree, Richardson, & Orphanos, 2009). Thus, there is a need for increased understanding of how PD supports teacher learning and instructional practice (Darling-Hammond et al., 2009; Education Week Research Center, 2014).

Heeding this need, in the past two decades, there has been increased attention to how different forms of PD can improve classroom instruction and student learning (e.g., Avalos, 2011; Garet, Porter, Desimone, Birman, & Yoon, 2001; Taylor & Colet, 2010). Many recommendations include PD in the form of collaborative teacher learning in teacher communities (e.g., Bryk, Lee, & Holland, 1993; DuFour & Mattos, 2013; Hadar & Brody, 2010; Louis & Marks, 1998; Skerrett, 2010; van Es, 2012). While there is consensus emerging on the factors of high quality PD, there are still many different approaches to school-based Teacher Learning Communities (TLCs).

Research suggests that one way to improve design, implementation, and evaluation of TLCs is through a deeper

* Corresponding author. College of Education, Michigan State University; 620 Farm Lane, 203 Erickson Hall, East Lansing, MI 48843, USA.

E-mail address: kintztar@msu.edu (T. Kintz).

¹ All of the second authors contributed equally to the manuscript.

understanding of the different configurations of TLCs that promote teacher learning (e.g., see Wayne, Yoon, Zhu, Cronen, & Garet, 2008). Vescio, Ross, and Adams (2008) found that well-developed TLCs had a positive impact on both teaching practice and student learning. In addition, research on the combination of learning conditions in TLCs can enhance understanding about how teachers learn best in local settings (Cobb, McCain, de Silva Lamberg, & Dean, 2003).

To address these needs, we adopted a configurational approach (Ragin, 2008; 2014) to examine factors that contribute to the types of talk that support teacher learning in TLCs (e.g., see Lord, 1994; Nelson, Slavit, Perkins, & Hathorn, 2008; Stoll, Bolam, McMahon, Wallace, & Thomas, 2006). We conducted a three-year, mixed method study of teachers involved in Formative Assessment for Michigan Educators (FAME), a PD program that was designed and initiated by the Michigan Department of Education (MDE) to support teachers' collaborative learning about formative assessment (FA). Because FAME was enacted through TLCs, it provided an ideal opportunity to examine factors that influenced discussions in a specific reform initiative designed to promote teacher learning about instructional practices.

2. Literature review

2.1. Formative assessment as knowledge and practice

Formative assessment (a.k.a. assessment for learning; Wiliam, 2011) is a teaching practice that informs both instruction and student learning (Bell & Cowie, 2001). Formative assessment is a process that uses students' ideas to guide both teaching and learning (CCSSO, 2008). It helps teachers be more reflective about students' understandings (Furtak, 2012) and more likely to support students in identifying barriers to learning (Marshall & Drummond, 2006). The use of FA can improve student learning (e.g., Black & Wiliam, 1998) and student involvement (e.g., Brookhart, 2013). However, learning about and enacting FA is complex for teachers, because they need to develop multifaceted knowledge and practice embedded in teaching and learning. Thus, professional development about FA has been proposed to help teachers learn and improve in this practice (e.g., Popham, 2009; Schneider & Randel, 2009).

2.2. Effective professional development and Teacher Learning Communities

Researchers have begun to illustrate connections among the design of PD, teachers' learning, and subsequent changes in classroom practice (e.g., Borko, 2004; Wilson, 2013). There is growing evidence that effective PD should address aspects of school capacity (King, 2002; Newmann, King, & Youngs, 2000); persist over a long period (Wei, Darling-Hammond, & Adamson, 2010); ensconce teachers in supportive communities of practice (Darling-Hammond, 1997; Stoll et al., 2006; Wenger, 1998); align with teacher, school, and district goals for student learning (Penuel, Fishman, Yamaguchi, & Gallagher, 2007); and engage teachers in authentic problems (Lave & Wenger, 1991; Webster-Wright, 2009; Wilson & Berne, 1999). Teacher learning communities are one form of PD that provide opportunities for learning by facilitating collaborative work (Grossman, Wineburg, & Woolworth, 2001) and promote collaborative inquiry (Nelson et al., 2008). Two assumptions justify the use of TLCs: (1) knowledge is situated in teachers' practical experiences and understood by critical and collective reflection, and (2) active participation in this process is related to increased knowledge and students' learning (Vescio et al., 2008).

Although TLCs have the potential to promote and sustain

learning, further evidence is needed to uncover why some TLCs are more successful than others. There is a large difference between a group of teachers sitting in a room for a meeting and teachers engaged in critical inquiry to further their professional growth. Some argue that effective TLCs benefit from a specific content focus and location (e.g., Slavit, Nelson, & Kennedy, 2010) and be led by experts (e.g., Stein, Smith, & Silver, 1999). Others suggest that inter-school TLCs can be effective (e.g., Richmond & Manokore, 2011) and that leadership should be distributed among the TLCs members (Stoll et al., 2006). Further research has found that the role of a coach was a critical factor in supporting a TLC (Grossman et al., 2001; Richmond & Manokore, 2011; Stein et al., 1999). In terms of composition, TLC diversity (accompanied with facilitation) can enrich discussion when problems of practice are raised and discussed from different perspectives (Slavit, Laurence, Kennedy & Holmund-Nelson, 2009; Thomas, Wineburg, Grossman, Myhre, & Woolworth, 1998).

2.3. Communities of inquiry

Levine (2010) clarified affordances of different conceptions of TLCs. Teacher communities have been referred to as, among others, instructional communities of practice (Lave & Wenger, 1991; Supovitz, 2002) and professional learning communities (DuFour, Eaker, & DuFour, 2005). For the purposes of our study, we will use the term "communities of inquiry" (e.g., Cochran-Smith & Lytle, 1992; Curry, 2008) to characterize the type of collaborative format and the nature of the teacher learning intended in the FAME PD program. Although the empirical realities of the cases in this study vary from this ideal, the construct of communities of inquiry (CIs) provides a conceptual and analytic lens to examine the nature of discussion within TLCs.

In CIs, teachers come together to problematize common practices and underlying assumptions, often through consideration of extant practices and student work. In formal instances, teachers participate in a "cycle of inquiry" in which "teacher researchers go through recursive stages of formulating problems, collecting data, analyzing data, reporting results, and planning for action" (Levine, 2010, p.112). The mechanism for learning embedded in CIs is straightforward—participating teachers learn through focused investigation and challenging of extant practices and beliefs. The importance of critical reflection in CIs has been illustrated in several studies. Wilson and Berne (1999) suggest that productive CIs must focus on building "trust and community while aiming for a professional discourse that includes and does not avoid critique" (p. 195). By challenging implicit assumptions and questioning axiomatic practices, CI discussions can lead to changes in practice (Antonacopoulou, 2004; Boud & Walker, 1998; Brockbank, McGill, & Beech, 2002; Brookfield, 2005; Katz, Sutherland, & Earl, 2005).

Similarly, Lord (1994) suggests that to support teachers in changing their practice, PD should allow teachers to ask questions about the practice of teaching and reflect on their own practice. However, just asking questions and reflection may not be enough to promote teacher change:

[f]or a broader transformation, collegiality will need to support a critical stance toward teaching. This means more than simply sharing ideas or supporting one's colleagues in the change process. It means confronting traditional practice – the teacher's own and that of his or her colleagues – with an eye toward wholesale revision (p. 192).

This idea of *critical colleagueship* (Lord, 1994) is an essential part of CIs that is often difficult to attain because these critical interactions may be against the personal and "experiential" nature of

the teaching profession (Labaree, 2003) and may run contrary to the culture of teaching where teachers are not often asked to explain their actions (Lachance & Confrey, 2003).

2.3.1. Discussion in communities of inquiry

Building of critical collegiality can be promoted through discussions in CI meetings. Lieberman and Mace (2009) found that when accomplished practitioners (i.e., those who looked deeply at practice, examined its complexity, and articulated what they have learned about their practice) made their teaching public, it provided the opportunity for teachers to learn about their own practice and involve a group of peers in the reflective inquiry. Similarly, Horn and Little (2009) found that conversational routines, specifically, the patterned and recurrent ways that conversations unfolded in two CIs, located in US urban schools, determined whether the conversation turned toward or away from teaching as an object of collective attention. Their findings indicate that resources and conversational routines in a CI contribute to the extent to which teachers discuss instructional practice and their subsequent learning regarding their practice.

In a study on teacher engagement in collaborative inquiry, Nelson et al. (2008) found that collaborative norms and using processes such as protocols to structure discussions promoted inquiry in CIs. In addition, the distribution of leadership responsibilities and construction of the inquiry focus helped the CI to develop and maintain collaborative inquiry in their discussions. The authors conclude that further research is needed on the conditions that contribute to sustaining effective CIs that facilitate teacher learning. Overall, these studies suggest that rich discussions on teaching and professional practice can be promoted in CIs, although support and structure needs to be provided.

There is great complexity and variability in studying the diverse conditions that support the discourse in a CI. Thus, a methodological approach is needed that examines multiple configurations of different conditions that may influence the nature and depth of discussion among CI members. The current study explores the necessary conditions for promoting in-depth and critical discussions in CI meetings. Specifically, we ask, what conditions are associated with greater depth of discussion in CI Meetings? The research question guided the procedures used to address this question including the selection of variables, analysis of data, and the discussion of the findings.

3. Methods

3.1. Study context

The CIs involved in our study were situated in various schools across the state of Michigan in the U.S. Each CI was composed of one coach (typically a teacher or administrator) and five to eight CI members. The FAME model allows for a great deal of local discretion in the formation of CIs and the structure, frequency, and specific content of the meetings themselves. The MDE provides some guidance and resources for the CIs including a single day, in-person meeting introducing the concept of FA as well as print and online resources. Through this design, FAME is intended to promote teacher learning about FA and increased use of FA strategies in the classroom. Since this learning process takes time, it is expected that CI members will commit to working together for three or more years (MDE, 2011). Each CI determines the scope and sequence of their FA study over the course of the year.

3.2. Participants

The research team consisted of two faculty members and a

combination of nine graduate students over the course of the study; four members of the research team are authors of this manuscript. In order to recruit a diverse set of participants, the director of the FAME program (from MDE) provided a list of CIs that ranged in their experience, school demographics, and student achievement outcomes. A member of the research team contacted the coaches of the CIs to ask if they were willing to participate in the study. Contact with the coaches and all aspects of the study were conducted in accordance with the university's institutional review board. All participants consented to have their meetings videotaped and all identities of participants and schools are protected using pseudonyms. Table 1 outlines the 17 CI cases included in the study.² As evident from the table, there is a range in make-up of CIs, as well as a range of years of participation.

3.3. Data sources

3.3.1. Video data

We videotaped CI meetings for twelve different CIs over a three-year time period from 2010 to 2013. One researcher attended each meeting, explained the research project, and obtained consent to videotape. The researcher observed all meetings and recorded video (as a non-participant observer). Table 2 shows the amount of video we gathered from each CI. In all, we examined over 74 h of video.

3.4. Constructs and outcome variable

Research suggests that much can be understood about teacher learning through an analysis of talk, such as discussions in CIs (Stoll et al., 2006). In order to characterize the discussion that happened in CIs, we used a recursive coding process, employing both top-down and bottom-up procedures. We began with modified grounded theory (Strauss & Corbin, 1990) to develop constructs that characterized the meetings. We initially used themes identified by the literature (e.g., active learning, collaborative inquiry, and reflective dialogue) and the PD materials (e.g., the way the content was described). We made several passes through the data to refine and clarify these initial categories to describe the various conditions on the CIs that may contribute to the depth of discussion.

We then developed rubrics based on these categories and examined the video of the CI meetings to elaborate on these predetermined categories using emergent themes that were suggested by a close study of teacher interactions in CI meetings. As we analyzed the data over the course of three years, the categories were further clarified and refined. The categories that became our initial focal constructs included: (1) FA content covered in the meetings, (2) role of the coach, (3) types of questioning (and by whom), (4) types of feedback (and by whom), (5) purpose of the team, (6) focus of the conversation on theory and/or practice, and (7) type of activity. The constructs were defined with examples to best capture the condition in the CI. We went through several iterations of the coding rubrics for the meetings as we continually refined the constructs. (See Appendix A for the video analysis protocol).

Although we characterized the meetings with several different constructs, we focus on the ones that had the most explanatory power

² Although we initially collected data on thirteen CIs, we omitted one CI, Fairgrove, from our study because we only had data for a single, 1-h meeting for this CI. The criterion for inclusion in our study was data for two or more meetings to ensure we had adequate data to identify patterns within and across CIs. Therefore, our study includes data for 12 CIs and a total of 17 different CI cases to account for the CIs that participated for more than one year of the study.

Table 1
Key characteristics of communities of inquiry.

Community of inquiry	Years	CI Make up	General content and structure of the meetings
Arvin	2010–11	2 Coaches: HS Spanish Teacher; curriculum coordinator CIM: HS teachers from different disciplines	Primarily examination of tool implementation, teachers' classroom videos and student work Secondarily discussion of book chapters and articles
Belmont	2010–11	Coaches: Curriculum and Instructional coaches CIM: 1 elementary group, 1 middle school group, 1 high school group	Primarily presentations by coaches Brief discussion of information from presentation
Bishop	2010–11	Coach: HS teacher CIM: HS cross-disciplines	Mixture of discussion of FA strategies and student evidence, sharing tools, and book study
Carson	2010–11	Coach: MS principal CIM: Elementary and MS teachers	Mixture of discussions of members' practices and examination of resources
Year 1			
Year 2	2011–12	Coach: MS principal CIM: Elementary and MS teachers	Mixture of working learning targets and associated tool and discussion of "student evidence"
Year 3	2012–13	Coach: MS principal CIM: Elementary and MS teachers	Mixture of reading and sharing examples from practice
Woodlake	2010–11	Coach: Elementary principal CIM: Elementary teachers in same building	Primarily discussion of book chapters, external constraints, unrelated topics
Maywood	2011–12	Coach: Elementary principal CIM: Elementary teachers	Primarily discussion of translating Common Core (National Governors Association, 2010) into student-friendly learning targets
Year 2	2012–13	Coach: Elementary principal CIM: Elementary teachers	Primarily discussion of gathering summative evidence of students and school closings in the district
Lincoln	2011–12	Coach: MS Principal CIM: Cross district, elementary, middle and high school teachers	Primarily reading articles and book and discussion of reading Discussion of development of a district "toolkit" of FA tools.
Year 1			
Year 2	2012–13	Coach: MS Principal CIM: Cross district, elementary, middle and high school teachers	Primarily reading and discussion of reading; off-topic discussion; continued development of the district toolkit
Stanton	2011–12	Coach: MS teacher CIM: Cross-disciplinary MS teachers	Mixture of reading a book; discussion of current FA practices; and designing student friendly learning targets and associated tools
Year 1			
Year 2	2012–13	Coach: MS teacher CIM: Cross-disciplinary MS teachers	Primarily discussion of Common Core State Standards and how they impact curriculum, instruction and assessment
Highland	2011–12	Coach: HS teacher CIM: Cross-disciplinary HS teachers	Mixture of designing FA tools and reflection on use of tools and discussion of external constraints
Albany	2011–12	Coach: Assistant principal CIM: HS mixed disciplines	Primarily planning a PD for all teachers in the school. Discussion of teachers who do FA well and who struggles
Roseville	2012–13	Coach: MS Teacher CIM: Cross-discipline Middle School	Primarily discussion about books and linking books to classroom practice
Bradbury	2012–13	Coach: English Teacher CIM: Cross-discipline Middle School	Primarily reading articles and finding tools (online) and discussion of tools

for our outcome variable, depth of discussion.

Outcome. We examined varying levels of discussion as the outcome for our study. *Depth of Discussion* refers to the extent to which CI members engage with the discussions, support each

other's ideas, and critically examine practice (i.e., see: [Lord, 1994](#); [Nelson et al., 2008](#); [Stoll et al., 2006](#)).

The depth of discussion in a meeting was coded into one of five categories: (1) presentations or other non-interactive activities which *did not include discussion*; (2) *one-way sharing* in which a CI member shared about an experience and then the conversation moved on to another topic; (3) *parallel sharing*, characterized by one CI member sharing an idea or example from practice, and then another CI member sharing an idea or example with little or no connection between the two topics; (4) *linking ideas or examples together*, when CI members built off one another's ideas; without pushing for reflection or in-depth analysis; (5) finally, *linking ideas and examples through a push for reflection or in-depth analysis*. In the segments of discussion that received a rating of 4 or 5, there needed to be at least one instance of the focal interaction involving three or more exchanges on a given topic. See [Table 3](#) below for the levels, description, and examples of depth of discussion.

Greater depth of discussion included more than simply linking ideas. Our conception of high depth of discussion includes dialogue in which CI members: provide critical feedback and in-depth analysis, challenge existing notions of practice, question the reasoning behind decisions, and demonstrate an openness to new ideas. Our study looked at combinations of conditions across CIs that shaped this type of in-depth discussion.

Table 2
Video data for each learning team.

Learning team	Meetings and hours of video
Arvin	2010–11: 3 × 1 h meetings
Belmont	2010–11: 2 × 1.5 h meetings
Bishop	2010–11: 2 × 2 h meetings
Carson	2010–11: 4 × 1 h meetings
	2011–12: 6 × 1–1.5 h meetings
	2012–13: 6 × 1.5 h meetings
Woodlake	2010–11: 2 × 1 h meetings
Maywood	2011–12: 7 × 1 h meetings
	2012–13: 6 × 1 h meetings
Lincoln	2011–12: 5 × 1 h meetings
	2012–13: 12 × 1 h meetings
Stanton	2011–12: 5 × 4 h meetings
	2012–13: 2 × 4 h meetings
Highland	2011–12: 5 × 1–1.5 h meetings
Albany	2011–12: 2 × 1.5 h meetings
Roseville	2012–13: 3 × 1 h meetings
Bradbury	2012–13: 4 × 1 h meetings

Table 3
Levels of depth of discussion, description, and example.

Level	Description	Example
1	No discussion	Coach: Now I would like to present the next section on peer assessment. Please refer to page 6 of your handouts.
2	One-way sharing	Teacher 1: I used a ticket out the door with my students to assess their learning ... " Coach: Moving on, let's read this article about formative assessment.
3	Parallel sharing	Teacher 1: In my class, I have learning targets posted on the wall. That way, all of the students can see them. Teacher 2: I struggle with not grading the students on formative tasks.
4	Linking ideas or examples together	Teacher 1: At the end of every unit, I have the students fill out test correction forms and I give them written feedback on their work. Teacher 2: I also have them work with a partner to correct their work and I conference with them about the written feedback to be sure they understand. Teacher 3: I would like to try that more in my class.
5	Linking ideas and examples through a push for reflection or in-depth analysis	Coach: What do you think contributed to your students' improvement? Teacher 1: I used this tool with my students and I saw such a difference in their scores. Teacher 2: I wonder if we used that tool in different classes whether we would see the same growth. I'm not sure it works the same way in math. Teacher 3: I have designed a similar protocol in math for my students and they responded well; however, some students feel challenged by a new approach to learning. You have to set them up for it.

3.5. Analytic techniques

The first step in reducing the data (Miles & Huberman, 1994) was to code each CI meeting video. We began the process by dividing the meeting video into codeable segments. Segments were defined by a change in the activity, content, or depth of discussion. We recorded the length of each segment and then assigned a code for each segments based on the constructs mentioned above. After an initial calibration process with the whole research team, two researchers coded each CI meeting video using a social moderation process and came to a consensus about the codes (e.g., Frederiksen, Sipusic, Sherin, & Wolfe, 1998). If a consensus could not be reached, the video was brought to the entire research team, who worked together to resolve the discrepancy.

Using these codes, we developed graphs that represented the percentage of time CIs spent in each code (or the percentage of probing questions CI members and coaches asked out of the total number of questions asked). This allowed us to reduce our data and create displays that could help illustrate the nature of each CI meeting and patterns within meetings over time (Miles & Huberman, 1994).

Using these graphs and our in-depth knowledge of the CIs, we wrote rich case studies of the CIs each year. This approach to case study research helped us to describe and explain phenomena that occurred in real (not researcher-altered) contexts (Stake, 2013). The case studies included a contextual description of each CI (e.g., demographic information about the school and district), the make-up of the CI (a description of members of the team and what they taught), a narrative of each CI meeting, description of each CI member's involvement in the discussions, and any relevant information helpful in the interpretation of the case (e.g., a school closure the day before a CI meeting). For the final analysis, we considered each year that the CI participated as a separate case to take into account the difference in CI composition and focus between years. This allowed us to analyze each CI case individually and to isolate which combination of conditions contribute to depth of discussion.

Each year, and again after our three years of data collection, we developed a descriptive matrix of each case that included a synthesis of the codes and narrative case features. This allowed us to make comparisons across cases to build explanations (a special type of pattern matching analysis) from the developing themes (Yin, 2009). We then examined the relationships among the conditions using fuzzy set Qualitative Comparative Analysis (fs/

QCA). We include a brief overview of fuzzy set QCA methods below.³

3.5.1. Qualitative comparative analysis

This study employs QCA, an analytic technique that identifies the minimum combination of conditions that contribute to a given outcome (Ragin, 1987, 2000, 2008). Grounded in set theory, this approach is ideal for examining the association among multiple complex conditions and a given outcome. Qualitative Comparative Analysis allows for significant levels of causal complexity (e.g., Fiss, 2007; Ragin, 2000, 2008) and accounts for the complexities of social reality by analyzing how one variable in combination with other variables is associated with the outcome (Ragin, 2008). Unlike regression analysis, in "conjunctural" causation, variables are not independent of one another and they do not compete with one another to explain variation in the outcome. Rather, the configurational approach forces researchers to consider how the effects of the values of variables work together to bring about an outcome. Furthermore, there may be different clusters of variables or conditions that can work together to achieve a result. That is, there may be multiple combinations of conditions and more than one way to realize an outcome.

Based on a conceptual model, QCA provides several advantages that allow researchers to identify the combination of conditions that are necessary and/or sufficient to produce an outcome. As demonstrated in other research, the configurational approach allows for cross-case comparison of cases (generally 8 to 150 cases) in which a certain outcome has occurred, compared to cases where it has not (e.g., Rihoux & Lobe, 2009; Sebastian, Allensworth, & Stevens, 2014; Trujillo & Woulfin, 2014). In addition, QCA considers the important role of context to understand relationships and how they connect to an outcome. The approach also helps to construct an explanation of how an outcome occurs by identifying cross-case commonalities and differences using both in-depth case oriented (qualitative) research as well as variable-oriented (quantitative) research. At the same time, some researchers have cautioned against the approach, noting the limitations of software to produce the solutions rather than theory (Cooper & Glaesser, 2011) and expressing skepticism about the difference from regression analysis (Achen, 2005). Nonetheless, this approach can further contribute to our

³ For a comprehensive description of the procedures specific to set-theoretic methods please refer to further literature on the topic (e.g., Fiss, 2007, 2008; Ragin & Fiss, 2008).

understanding of the conditions of teacher collaborative learning that promote high depth of discussion.

3.5.2. Employing qualitative comparative analysis

We used the 17 cases described above to examine conditions associated with depth of discussion. Drawing on theory and in-depth knowledge of these cases, we determined the potentially necessary and sufficient conditions for the outcome. As previously described, we coded each CI meeting for evidence of seven potential conditions (FA content, role of the coach, questioning, feedback, purpose of the team, focus of the conversation in theory and/or practice, and type of activity), and five different levels of the outcome, depth of discussion (no discussion, one-way sharing, parallel sharing, linked ideas and examples with no push for reflection, linked ideas and examples through a push for reflection or in-depth analysis).

Next, because we used a variation of QCA, fuzzy set analysis (fs/QCA) (Ragin, 2000, 2008) based on Boolean algebraic techniques in which variables are assigned values ranging from 0 (absent) to 1 (present), we assigned qualitatively-anchored values to each condition and the outcome. To assign values, we first determined thresholds for case membership ranging from 0, indicating a completely absent condition, and increasing in 0.2 increments to 1, indicating characteristics of a completely present condition for each CI. In other words, the values represented a gradation of different degrees to which the conditions occurred in the context of CI meetings. Next, we assigned each case the values from 0 (absent) to 1 (present) for each of the conditions (i.e. variables) included in the fuzzy set. Unlike traditional quantitative measures (e.g. ordinal variables) the fuzzy-set membership values in this study have “qualitative anchors” that describe the characteristic meaning of each assigned value. Assignment of the fuzzy set scores was based on our comprehensive analysis of the qualitative data. (See Appendix B for the criteria used to determine thresholds for case membership). Then, we constructed a matrix with all of the values from 0 to 1 for each condition and the outcome corresponding to each case.

Next, to analyze the causal complexity and association among the cases using fs/QCA, we constructed a truth table (Ragin, 2008; 2014) that included each condition and the outcome with all corresponding combinations of values. Through the use of Boolean algebra, the truth table indicated the combination of conditions that were associated with the outcome. Finally, we generated a list of expressions, known as a *solution*, which showed the combinations of conditions linked to the outcome. This included analyses for both coverage and consistency. Coverage determines empirical relevance, i.e., the extent to which we observed a causal combination in relation to the total number of cases in our sample. Consistency is the degree to which the cases include a condition that leads to the outcome (sufficiency) or one of multiple causal conditions (necessity).⁴

4. Findings

In this section, we discuss the overall findings for depth of discussion for each CI. We then delineate three primary conditions related to the depth of discussion in CIs that emerged from our

analysis: the purpose or intended objectives of the CI, coach probing questions, and the connection of theory to practice. Finally, we present the fs/QCA findings and the interplay of the conditions associated with the outcome.

4.1. Depth of discussion

Overall, we found a high level of variation in the depth of discussion among the CIs, as represented in Fig. 1. The discussions in the CI with the lowest depth of discussion, Maywood 2, consisted primarily of one-way and parallel sharing. In this CI, members would share about using learning targets in their classroom individually without building on each other's ideas or pushing for reflection. For example, one CI member described the use of science learning targets in the following way, “In my class, I refer to the learning targets throughout the lesson.” This was followed by an unrelated description of student work in mathematics from another teacher, “I usually expect my students to turn in their work at the end of the period, after we have gone over the assignment.” This example is representative of the parallel sharing that occurred in discussions at the left side of the figure.

On the other hand, the CI with the highest depth of discussion, Stanton, participated in discussion that linked ideas together as well as discussion that linked ideas together with a push for reflection. For example, in one instance, the Stanton CI members built FA tasks and CI members pushed each other to be reflective about how they would be implemented in the classroom and how to ensure students demonstrated mastery of standards. The following transcript and notes demonstrate the discussion that linked ideas together and pushed for reflection (See Table 4):

In this discussion, the CI members offered different suggestions in response to one members' question about checking for mastery in a formative way. Other members also engaged in comments where they raised questions and presented counter viewpoints or contrasting ideas. The discussion provided the members with an opportunity to examine a particular aspect from their instruction in light of FA. Other in-depth discussions demonstrated a similar nature of inquiry and collective reflection. Overall, CI members engaged in more critical dialogue and reflection toward the right side of the graph.

Given this variation, we examined the conditions that were present in the cases that had high and low depth of discussion to understand factors that contributed to this outcome. Three main conditions emerged from the cross-case analysis: purpose of the CI; coach questioning, and connecting theory to practice. Below we illustrate how each of these factors is related to the depth of discussion.

4.2. Three primary conditions

4.2.1. Purpose of the learning team

The purpose condition is defined by the explicit goals and objectives for each CI. CIs tended to have either a single purpose or a dual purpose. For example, the single purpose CIs focused solely on the topic of FA. The agenda for these CI meetings emphasized goals and objectives relevant to learning and sharing about FA. In contrast, dual-purpose CIs combined the goals and objectives of other initiatives into the FA meeting agenda. These CIs also performed the responsibilities of another initiative such as school improvement, *Reading Recovery* or *transition to common core state standards* teams.⁵

⁵ We captured the information about the purpose of the CI based on the activities and objectives of the meeting.

⁴ Since we were able to quickly determine that if any causal combination (necessity) relationship existed in our data it would include the outcome, we used the following formula: Consistency ($Y_i \leq X_i = \Sigma[\min(X_i, Y_i)] / \Sigma(Y_i)$) Where the consistency of membership in the outcome is determined by the minimum of X (membership in the causal combination) and Y (membership in the outcome) or the *i*th case divided by the sum of membership in the outcome. This number is then divided by the total number of cases in the configuration, yielding a consistency score between 0 and 1.

Fig. 1. Depth of discussion for all CI Cases.

Table 4
Example depth of discussion code 5 transcript and description.

Transcript	Notes
Teacher 1: I also have a question about how we can incorporate FA into class A ... since we have the objectives for science, and we have the ability to look at, and have them work for mastery, and they have a way there They have all that data available and I'm wondering how we can use that, formatively. Does it make sense? So reflect on how we can know if we are reaching and meeting the content standards?	Teacher 1 asks a probing question about implementation and student mastery of standards
Coach: Can someone print out a list? Is there some 'I can statements', or make those 'I can statements?' Do we have that?	Coach requests action in response to the question
Teacher 1: I have done that!	One CI member has already done that
Teacher 2: We have	Second CI member responds as well
Teacher 3: It still has to have an objective task not just a list. You wouldn't have to put necessarily in there a grade, but it has to be designed like an objective task to give them the skill.	Teacher 3 pushes back on what type of task would be appropriate, connecting ideas to the previous comments
Coach: If you can just, put the objective and even include the 'I can statements' in front of that, and they know what they matched to, they could self-assess that way. And then we'll have all the graphs and the data, even if they just do "yes/no", or how are we going ...	Coach responds extending the topic of conversation and building on the previous comment.
Teacher 1: And having the kids assess themselves: "I strongly believe that I can determine the number electrons by looking at the periodic table." But then, do we have them design "show me"?.. Do you know what I'm saying?.. So they can say, "yes I can ..."	Teacher 1 connects to previous idea and further reflects on instructional practice regarding student self-assessment and whether the student has mastered the standards. Asks a question.
Coach: Well, in that case I have to come back after they take the summative task and say, hold on! 80% of you told me that you could do this, but only 20% of you could do it. There is supposed to be some trust there too ... You are not going to be graded on it, and may even not seeing an A or things like that, you should be as honest as you can ... 'cause it is the same thing to get an exit slip, or hold the green card, the red card or whatever.	Coach responds by extending the conversation and explains the approach to connect student self-assessment responses with performance on summative task to determine whether students are reaching and meeting the standards.
Teacher 2: Even just the reflection formatively, you know, would be good, because, they can say, "I thought I understood this, but maybe it is the wording", they can specifically say, "the specific wording on this question lead me to believe this" Whatever, they can actually look at the individual questions before the test	Teacher 2 further builds on the conversation by emphasizing that student reflection on their self-assessment compared to their performance on a summative task could help to clarify their understanding of the material.

When a CI had a single purpose focused on FA, they were able to spend more time discussing FA related theory and practice; however, it was not meeting time focused on FA alone that contributed to depth of discussion. The Belmont CI spent a considerable amount of time focused on FA, but because it was a dual purpose team, the meeting focused on presenting large amounts of material with few opportunities for substantive discussion. Fig. 2 shows the depth of discussion for the single purpose CIs, arranged from the least to greatest depth of discussion. All of the single purpose CIs engaged

in discussion that linked ideas and examples with a push for reflection, with a range of just over 10% to over 75% of the conversations involved linking ideas with or without reflection.

Alternatively, dual-purpose CIs engaged in less discussion that linked ideas with a push for reflection (see Fig. 3). Four of the seven CIs did not engage in any discussion that linked ideas or examples with a push for reflection. The three CIs that did engage in discussion that linked ideas or examples with a push for reflection did so less than 10% of the time, and the total time of linking ideas or

Fig. 2. Single purpose teams and depth of discussion.

examples with and without a push for reflection did not exceed 30% of the total meeting time. These CIs had other goals and objectives that competed with or detracted from their in-depth discussion of FA.

At the same time, a few of the single and dual-purpose CIs appear to be exceptions to the rule. Single purpose CIs Carson 1 and Bishop both had less than 3% of the discussion that linked ideas and examples with a push for reflection. Although these CIs had a single purpose, it did not guarantee a high depth of discussion. Similarly, dual-purpose CIs Lincoln 1 and 2 had over 5% of the discussion time linking ideas and examples with a push for reflection, and almost 30% of the discussion involved linking ideas and examples with or without reflection. The dual purposes of these CIs did not prevent them from engaging in reflective, deep discussions. These examples suggest that other conditions, beyond the purpose of the CI, also contribute to the depth of discussion of the CI.

4.2.2. Coach questioning

The actions of the coach in the CI discussion contributed to conversations in the CI meetings.⁶ Coaches who facilitated the conversation without overly guiding it and mediated CI members thinking by asking probing questions tended to promote high depth of discussion.⁷ One example of this is when the coach of Carson 2 set up several tools on chart paper and used a gallery walk protocol to have CI members discuss new FA tools. She then asked probing questions to have the CI members think about relating the tools to their classroom practice. The coach prompted the CI members to be reflective and the protocol supported them to connect their ideas.

⁶ While there are many ways a coach may interact with a CI, our qualitative analysis focused on two observable coaching behaviors that were associated with CI conversations (1) questions, and (2) feedback. Specifically, the coach behavior included the extent to which the coach asked probing questions and provided feedback to move practice forward. In this section, we focus on questioning because it provided more explanatory power for depth of discussion (see QCA findings).

⁷ We identified two main types of questions associated with the depth of discussion in CIs, clarifying questions and probing questions based on the Cognitive Coaching model by Costa and Garmston (1994). In this model, coaches learn to mediate participants thinking by asking a range of questions to promote reflection. Clarifying questions serve to gather information about a topic, such as, *What subject do you teach?* Probing questions provide opportunities for reflection, contemplation, synthesis, that require individuals to make connections and explain their thinking. For example, a coach may ask, *Could you tell me more about how the way you used that strategy was different from the science teacher?*

As a result, the CI members engaged in critical inquiry and deep discussion.

Alternatively, in the Belmont CI, the coach presented information about FA and tools without providing opportunities the CI members to discuss or reflect on these tools. The coach performed the role of an expert and conducted workshops in which the participants were recipients of the information without engaging in probing questions regarding their practice. The members of this CI had minimal time to meet in groups and engage in collaborative inquiry. In turn, there was little discussion among the CI members, and no deep discussion involving reflection. Fig. 4 depicts the percent of probing questions compared to the total number of questions asked by the coach.

In Fig. 4, the CIs are arranged from least to greatest depth of discussion. Although there is variability throughout the graph, communities on the right side of the graph, i.e., those with greater depth of discussion, tend to have more probing questions. In contrast, there were fewer probing questions in communities with lower depth of discussion, for example, in the Stanton 2 CI, only about 10% of the questions initiated by the coach were probing questions. The CIs with a lower percent of probing questions tended to focus more on clarifying questions or gathering information questions, such as in Stanton 2 the coach asked questions like, *How many students are in your class?* These types of questions did not promote linking ideas or a push for reflection. On the higher end, in the Carson 2 CI over 40% of the questions initiated by the coach were probing questions. The coach asked questions such as, *What are some ways that you have used this tool to promote student understanding?* The probing questions provided opportunities for CI members to reflect on their practice and FA theory while at the same time extending on other CI members' ideas. While there were some patterns in questioning, we see variability throughout the graph, leading us to conclude that there may be other contributing factors in explaining the Depth of Discussion.

4.2.3. Content or structure that promoted connection of theory and practice

The final condition associated with a high depth of discussion refers to the attention given to theory, and the ability to then apply the theory to practice in addressing authentic teaching problems (Wilson & Berne, 1999). We found that the extent to which the content promoted an opportunity to connect theory and practice ranged among five levels, (1) unrelated content, i.e., conversations

Fig. 3. Dual-purpose communities of inquiry and depth of discussion.

that did not address FA (e.g., discussion of recent school events); (2) abstract discussion of the content, i.e., conversations about how individuals should know more about FA practice, without relating to theory or practice; (3) discussion of theory only, for example discussing a book, but not linking the ideas to classrooms, (4) sharing of practice only, for example sharing artifacts or stories without linking to the underpinnings of FA, and (5) discussions linking theory and practice such as considering how FA theory was linked to specific instruction in the classroom.

Fig. 5 illustrates the range in how CIs made connections between theory and practice in their discussions. In Highland, Maywood 1, Carson 2, and Stanton 1, the CI members made more connections between theory and practice. For example, in the Highland CI the teachers made connections between the theory of motivation and providing written feedback to students in the classroom so they could track their progress. When CIs focused on a

specific classroom practice or pieces of student work (as opposed to abstract discussion of FA techniques), colleagues tended to be able to delve more deeply into their ideas. For example, in the Maywood 1 CI, comprised of K-4 teachers, the most critical conversations took place around one teacher's presentation of a specific learning target and linking that to the importance of goal setting with students. For the most part, the CIs with higher depth of discussion (right side of the graph) also made connections between the FA theory and classroom practice that provided opportunities to engage in reflective dialogue and critical collegiality.

Alternatively, the discussions in three other CIs, Stanton 2, Belmont, and Woodlake, never linked theory and practice. Stanton 2 CI members focused solely on sharing ideas about a new reform initiative at the school without connecting FA theory with instructional examples. The teachers spent a lot of time sharing experiences from their classrooms with little connection to theory

Fig. 4. Percent of probing questions compared to the total number of questions (Note: CIs are arranged from lowest depth of discussion to highest depth of discussion).

Fig. 5. Connecting theory in practice.

that could be used to challenge or move their practice forward. In another example, Belmont primarily participated in presentations facilitated by instructional coaches. The presentations provided information on FA theory, with little opportunity for the teachers to engage in discussion about their teaching in relation to the theory.

4.3. Synthesis of all factors: necessary and sufficient conditions

Using findings from our case study analysis, we used *fs/QCA* to examine the necessary and sufficient conditions for high depth of discussion. A *necessary condition* is one that must be present when the outcome, depth of discussion, occurs; however, at the same time, the necessary condition does not guarantee the outcome (Ragin, 2000; 2008). That is, a necessary condition must be present for the outcome to be achieved; however, it does not achieve the outcome on its own. A *sufficient condition* is a condition, that when present, guarantees that the outcome, depth of discussion, will occur; however, it may be one of many conditions that produce the outcome (Ragin, 2000; 2008). For the purposes of reporting the *fs/QCA* data, we include only the conditions that were associated with the outcome. The other conditions listed above were not found to be necessary or sufficient to produce the outcome (i.e., they were not found to have any explanatory power). There were three conditions associated with the outcome: (1) single purpose, (2) coach probing questions, and (3) connecting theory and practice.

Table 5
Configurations related to high depth of discussion.

Configuration		Solution	
		1	2
Objectives	Single Purpose	•	•
Coach Questioning	Coach Probing Questions	•	⊗
Content and Structure	Connection of Theory and Practice	⊗	•
Consistency		0.67	1.00
Coverage		0.35	0.12

Note: Black circles indicate the presence of a condition, and circles with "x" indicate its absence.

Tables 5 and 6 illustrate the findings of our *fs/QCA* analysis. We use the notation developed by Ragin and Fiss (2008) in which dark circles ("•") indicate the presence of a condition, and a circle with a cross ("⊗") indicates the absence of a condition.

Table 5 below illustrates two solutions for high depth of discussion. The two solutions indicate there were two configurations that were associated with high depth of discussion. In Solution 1, CIs with high levels of the conditions single purpose and coach probing questions present and with low levels of the condition connecting theory and practice were associated with high depth of discussion. The total coverage for solution 1 was 0.35, meaning that 6 of our 17 cases exhibited this configuration. The first solution also has an overall moderate consistency of 0.67. The consistency showed that 67 percent of the time when the solution 1 configuration was present, it led to the outcome high depth of discussion. This consistency approaches the typical threshold used but does not exceed it (≥ 0.75 , Ragin, 2008); nonetheless it is a notable finding. In the second solution, CIs with high levels of the conditions single purpose and connecting theory and practice present, and with low levels of coach probing questions were associated with high depth of discussion. The coverage for solution 2 is rather low, it only occurs in 2 of our 17 cases. Yet, the consistency of 100 percent, is well above acceptable (≥ 0.75). This means, every time the configuration was present for those CI cases, it was associated with high depth of discussion. The finding of two overall solutions listed in Table 5 indicates a situation of across-case equifinality of solutions. In other words, there was more than one configuration that was associated with the

Table 6
Configurations related to low or non-existent depth of discussion.

Configuration		Solution		
		1	2	3
Objectives	Single Purpose	⊗	⊗	•
Coach Questioning	Coach Probing Questions	•	⊗	⊗
Content and Structure	Connection of Theory and Practice	⊗	⊗	⊗
Consistency		1.00	1.00	1.00
Coverage		0.12	0.24	0.18

Note: Black circles indicate the presence of a condition, and circles with "x" indicate its absence.

outcome of high depth of discussion.

In looking at solutions 1 and 2 (Table 5), the condition single purpose had to occur with either the condition coach probing questions or connecting theory to practice for a CI to achieve a high depth of discussion. Thus, single purpose was a necessary condition for high depth of discussion to occur. In our analysis, single purpose did not bring about a high depth of discussion on its own; indeed, only when single purpose was paired with high levels of the conditions coach probing questions or connecting theory and practice did it lead to high depth of discussion. We did not have any cases in our data that were high in all three conditions, but we hypothesize that CIs demonstrating high levels of all three conditions would also be associated with high depth of discussion.

Table 6 shows the three solutions for low depth of discussion. In 54 percent of our cases there was a low depth of discussion and these cases support the findings for high depth of discussion from Table 5 above. The three solutions indicate there were three different configurations which seemed to influence low depth of discussion. In Solution 1, CIs with the conditions coach probing questions present *and* the absence of single purpose and connecting theory and practice conditions were associated with low depth of discussion. The total coverage for solution 1 was 0.12, that is, 2 of our 17 cases exhibited this configuration. The consistency showed that 100 percent of the time the solution 1 configuration led to the outcome low depth of discussion. In the second solution, CIs with all three conditions absent, single purpose, connecting theory and practice, and coach probing questions, were associated with low depth of discussion. The coverage for solution 2 is 0.24, meaning it occurred in 4 of our 17 cases. The consistency was 100 percent, meaning every time the configuration was present (i.e., none of the three conditions were present) for those CI cases, there was low depth of discussion.

Solution 3 indicates that CIs with single purpose condition present, but with coach probing questions and connecting theory and practice conditions absent were associated with low depth of discussion. The coverage for solution 3 is 0.18, or 3 of our 17 cases exhibited this configuration. The consistency for this solution was also 100 percent; all of the cases with this configuration had low depth of discussion. Solution 3 supports our argument that single purpose is necessary to produce depth of discussion, but not sufficient on its own for a high depth of discussion. The presence of CIs with a single purpose that do not have a high depth of discussion does not undermine the assertion that single purpose is a necessary condition; those CIs may simply lack other conditions that need to be present with a single purpose for CI members to engage in high depth of discussion. This is in line with the understanding from the case studies in which single purpose alone did not lead to depth of discussion; rather single purpose may promote other factors such as the clarity for a coach to focus on certain questions, and the potential for CI members to make connections between a particular theory and their practice.

The finding of three overall solutions listed in Table 6 indicates a situation of across-case equifinality of solutions for low depth of discussion as well. In other words, there were three configurations that were associated with the outcome of low depth of discussion. Table 6 provides further evidence that the single purpose condition was necessary for high depth of discussion to be present in a CI.

Overall, the main findings indicate that single purpose was a necessary condition for the outcome to be present, and that all CIs in our study with a high depth of discussion had a single purpose. Given that single purpose is a necessary condition, it must be

present in all configurations that lead to the outcome. However, the single purpose was not sufficient to be associated with depth of discussion on its own.

In addition when coach probing questions or connecting theory with practice were linked with single purpose, the configuration presented a sufficient condition for high depth of discussion. Although we found that single purpose was a necessary condition for the depth of discussion outcome, this condition may not be enough on its own. This makes sense given that even if a CI has a single objective focused on formative assessment, the depth of discussion is not guaranteed to be high. Overall, when the explanatory condition single purpose was present together with coach probing questions or connecting theory and practice in any of our cases, they were both necessary and sufficient conditions for high depth of discussion.

5. Discussion

5.1. Combinations of configurations

Fostering cultures in which teachers in CIs can push on each others' ideas and offer professional critique is difficult (Dobie & Anderson, 2015; van Es, 2012); however when CIs are able to engage in these deep discussions, there is room for reflection and improved practice (Dobie & Anderson, 2015; Hadar & Brody, 2010; Stoll et al., 2006). Thus discovering specific combinations of conditions that promote deep discussions is critically important. In our study, we found that there was more than one way to promote a high depth of discussion. Our results indicate that having CIs focusing on a single purpose was a necessary condition for the CIs to engage in deep and critical discussions. Although having a single purpose in a CI was a necessary condition for high depth of discussion, this alone was not sufficient. Instead, in addition to the prior condition, for some CIs, having a coach able to ask high-level probing questions was needed (a sufficient condition) for high depth discussions. Other cases demonstrated that in addition to single purpose, being able to connect the theory of FA with their classroom practice was needed (a sufficient condition) for high depth discussions.

5.1.1. Practical applications

In the current educational context, schools and teachers are often inundated with a myriad of reform initiatives (Cooper & Shear, 2012). The finding from this study that a single purpose CI was associated with high depth of discussion suggests that schools aiming to promote teacher learning would benefit from a focused, coherent PD program that limits CIs to a single purpose. The opportunity to focus on one initiative over time could facilitate in-depth conversations involving critical inquiry and reflective dialogue. A single purpose CI may be more likely to attend to other known aspects of high quality PD such as basing PD reform on empirical evidence from teachers' experiences (Fishman, Marx, Best, & Tal, 2003) and providing opportunities for focused, collaborative work (Avalos, 2011). This finding extends on previous research indicating the importance of instructional program coherence in educational reform (King & Newmann, 2001; Newmann, Smith, Allensworth, & Bryk, 2001). The current struggle for coherence and control amidst the United States federal government's shifting involvement in education policy reform (Grissom & Herrington, 2012), further suggests the need for local educational contexts to provide teacher collaborative learning PD with a single purpose.

In addition, a successful CI has to be able to generate a group dynamic that enhances professional learning. A CI needs to recognize needs and promote teachers' initiative and leadership with regard to professional practice (Wei et al., 2010). A crucial aspect to promote teachers' professional inquiry and learning is the role of the CI coach. Teachers arrive to CI work with a set of diverse and individual expectations, and the role of the coach (or facilitator) is to converge these expectations through the implementation of common means and norms (Dooner, Mandzuk, & Clifton, 2008). In this study, the type of questions asked by the coach was associated with the depth of discussion in the CI. It is unclear whether the construct of probing questions is linked to other types of leadership qualities of the coach, however, the importance of the coach and her ability to ask high quality questions has implications for coach training and preparation in working with CIs.

Furthermore, in order for PD to be effective, teachers must be positioned at the center of their own learning. Engaging teachers in activities of understanding and facilitating their students' learning tends to lead to improvement in teachers' future practice (Franke, Carpenter, Feinman-Nemser, Ansell, & Behrend, 1998; Lotter, Harwood, & Bonner, 2006). In addition, sharing knowledge and teaching experiences within CIs is also a crucial component for improvement of teaching practice (Lieberman, 1995; Loucks-Horsely et al., 1998). CIs that address content and practices provide an opportunity to promote teacher growth around FA practices (Wylie, Lyon, & Goe, 2009). Nonetheless, change in teacher knowledge and practice takes time and focused effort. When CIs connected theory and practice, teachers were able to contribute their knowledge from the classroom, and make connections to new frameworks and ideas.

5.1.2. Theoretical implications: critical collegueship in communities of inquiry

This study of CI-based PD contributes to the theoretical understanding of the factors that promote critical inquiry and reflection among teachers. The focus on depth of discussion in this study provided an initial opportunity to capture some of the elements that may contribute to *critical collegueship* (Lord, 1994). Together, PD and critical inquiry in the form of CIs could serve as a useful resource for teachers seeking to engage in more ambitious practices. With this possibility, the specific conditions within CIs including the single purpose, coach probing questions, and connection of theory and practice that were associated with high depth of discussion also seem to be the factors that promote elements of critical collegueship. Future studies can extend upon these findings to further operationalize critical collegueship and uncover additional ways to promote such critical inquiry and reflection among teachers.

5.2. Advantages of the QCA/Configurational approach

The use of the *fs*/QCA methodology was a useful approach to identify commonalities across a relatively small number of cases. The central analytic strategy included examining cases sharing a given outcome (i.e., depth of discussion) and attempting to identify their shared causal conditions (e.g., the possibility that any or all of three conditions were present). As opposed to regression analysis, where conditions compete with each other to explain the outcome, *fs*/QCA allowed us to examine conditions and combinations of conditions that seemed to influence a high depth of discussion in CIs. Furthermore, *fs*/QCA allowed us to examine conditional patterns associated with an outcome

without losing focus on the cases themselves, what Ragin (2008) calls both case-oriented and variable-oriented research. Our finding that the conditions of single purpose, coach probing questions, and connecting theory and practice work together to promote depth of discussion has further implications for work in educational policy.

These findings suggest the need for a learning paradigm in PD policy that allows for multiple conditions to work together to achieve a given outcome. In this way, schools and teachers are able to have both the support and the flexibility to promote learning and implementation in local contexts. Further research will benefit from using methods such as QCA in exploring the interplay of these conditions and others in different educational settings.

5.3. Limitations and future directions

We are unable to claim generalizability to different contexts and TLCs with different combinations of conditions. Although our findings are potentially applicable to other reform initiatives in different settings, this study takes place in the context of one particular reform initiative in the U.S. In addition, the CIs were not organized by specific content area or grade level; therefore, we were unable to address the association of these factors, which we anticipate are important, with the depth of discussion in the CI.

Future research directions include an examination of necessary and sufficient combinations of conditions in different contexts and with different PD initiatives. Additional empirical study could focus on the implementation of different PD reform initiatives involving CIs at the elementary, middle, and high school levels, and CIs representing specific content areas. In addition, it will be beneficial to direct further research and resources toward opportunities for coaches that develop skills in facilitating critical inquiry and reflective dialogue among teachers. At the same time, promoting coaches' competence in posing questions and developing probing questions could support teachers in their efforts to improve instructional practice.

6. Conclusion

As U.S. educational policy increases expectations for student outcomes, teachers and their instructional practices are ultimately responsible for the success of these reforms. A more comprehensive understanding of teacher learning experiences can support teacher growth and development to improve instructional practice resulting in improved student outcomes. To achieve such ambitious policy demands for improving all students' learning, PD must address several aspects of school capacity to improve instructional practice (King, 2002; Newmann, King, & Youngs, 2000; Youngs & King, 2002). However, teaching is complex and policy has traditionally been a blunt instrument to effect lasting change (Cuban, 2013; McLaughlin, 1991, 2005). Fortunately, teachers do not mediate policy alone. Teachers can benefit from membership in supportive learning communities that can help them make sense of new ideas, examine their existing practice, and enact new instructional practices (Snow-Geron, 2005; Vescio et al., 2008). This study found that a single purpose for the CI focused on a clear objective, the role of coach asking probing questions, and meeting content or structure that allowed teachers to make connections between theory and practice were associated with in-depth discussions in CI meetings. Further research on reform initiatives involving CI-based PD can shed light on the factors associated with teacher learning and enactment of ambitious teaching practices. This documentation of the characteristics of effective CIs and the

factors that influence the shift from a group of teachers to “critical colleagues” over time provides information that can scaffold all CIs in promoting teacher learning, and in turn, improved instructional practice.

Conflict of interest

The authors state that there has been no conflict of interest in preparing this manuscript.

Funding

The research reported here was supported in part by the Michigan Department of Education and Michigan State University. The opinions expressed are those of the authors and do not represent views of the Michigan Department of Education or Michigan State University. The funding source had no involvement

in the study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

Acknowledgments

An earlier version of this paper was presented at the Annual Meeting of the American Educational Research Association, April 2015. The authors would like to thank the other members who participated in the research for this project: Joanne Philhower, Steve Bennett, Amanda Frasier, and Dawnmarie Ezzo. We would like to thank Ed Roeber, Kim Young, and the Michigan Department of Education as well as Michigan State University and the Educational Policy Center. We also appreciate Peter Youngs for his feedback on an earlier draft of this manuscript. The authors would also like to thank the teachers and principals who generously participated in this research.

Appendix A. FAME Video Analysis Protocol

Analysis of videos is generally guided by our research questions; however, we are more broadly interested in the types of interactions that occur in each type of learning team. Our research questions are:

Which types of learning teams are more effective^a?

- Content specific
- Grade-level specific
- School-based, multi-school ...

What happens in a learning team when you have a certain type of coach?

- Coaches' primary job responsibility
- Coaches' prior coaching experiences
- Coaches' prior knowledge of FA
- Use more FA tools/components
- Use FA tools more often
 - Planning for FA process
 - Learning targets in student friendly use given to students ...
- Effects student learning and achievement (teachers' perception)

We will analyze the video data in two or more rounds. The first round will be done at a large grain size.

We will record these analyses in a table and use them to identify what sections of video we want to do a more fine-grained analysis and, perhaps, transcribe.

1. Watching the video, identify large “chunks” or sections where a single discussion or type of interaction is taking place.

- a. In the table give an overview of the section. Include:

- Time stamp (beginning time, ending time)
- Who is talking (we can identify people as A, B, C ..., knowing who the coach is will be important)
- The content of the discussion
- Specific language in the discussion

2. Also include whether there is any evidence for the following:

- Teacher knowledge and practice around formative assessment
- Impact on student knowledge and practice
- Team building
- Organization
- Use of protocols
- Effective organizational strategies
- Feedback
- Use of resources
- Other interesting ideas that we may want to pursue
- After watching the video, write a summary (2–3 paragraphs) giving:

3. Key themes, patterns, and examples related to:

- Teacher knowledge and practice around formative assessment
- Impact on student knowledge and practice
- Role of the coach (or other LTM)
- Team building
- Organization
- Use of protocols
- Organizational strategies
- Feedback
- Appropriate use of resources

4. Any other ideas or lenses that you think that we can bring to these data. Include references that could help us think about further analyses.
-

^a Effective – defined as ability to help learning team members gain **knowledge** of and use **practices** around the formative assessment process.

Appendix B. Criteria for Degrees of Case Membership and Threshold Values

	Purpose	Role of coach	Content	Depth of discussion
Criteria and Degrees of Case Membership	0: Dual purpose	0: The team is completely out of the	0: The team is completely out of the Proximity to practice case. These teams did not spend time in discussion that involved linking theory and practice. The way we define completely out of the case is when the team spent 0% of the meeting time in discussion that linked theory and practice.	0: The team is completely out of the depth of discussion case. These teams did not spend time in discussion that linked ideas and examples through an examination of WHY things are similar/different or why they happened. The way we define completely out of the case is when a team spent 0% of the overall meeting linking ideas and examples through an examination of WHY things are similar/different or why they happened.
		0.2: Very Low questioning , less than 15% of the total percentage of questions include probing/reflective questions. Low feedback , less than 15% of the total percentage of feedback includes feedback with information to move practice forward.	0.2: Very Low proximity to practice : Less than 15% of activity code on discussion of potential uses of FA for student learning, teacher collaboration, or school-wide reform. (And/Or) Less than 15% of depth of content discussed involving linking theory and practice.	0.2: Very Low DoD , These teams are barely in the depth of discussion case because only a very small amount of the meeting time is spent involving discussion that: links ideas and examples through an examination of WHY things are similar/different or why they happened. We define a very small amount by the DoD codes that are 2–3% of the average of the overall meeting time involving discussion that: links ideas and examples through an examination of WHY things are similar/different or why they happened.
		0.4: Low questioning , 15–25% of meeting codes include probing/reflective questions. Low feedback , 16–25% of meeting codes include feedback with information to move practice forward.	0.4: Low proximity to practice : 15–25% of activity code on discussion of potential uses of FA for student learning, teacher collaboration, or school-wide reform. (And/Or) At least 16–25% of depth of content discussed involving linking theory and practice.	0.4: Low DoD : More out than in: Graph indicates 4–5% of overall meeting time involving discussion that: links ideas and examples through an examination of WHY things are similar/different or why they happened.
		0.6: Med questioning , 26–35% of meeting codes include probing/reflective questions. Med feedback , 26–35% of meeting codes include feedback with information to move practice forward.	0.6: Med proximity to practice : 26–35% of activity code on discussion of potential uses of FA for student learning, teacher collaboration, or school-wide reform. (And/Or) 26–35% of depth of content discussed involving linking theory and practice.	0.6: Med DoD : More in than out: Graph indicates 6–10% of overall meeting time involving discussion that: links ideas and examples through an examination of WHY things are similar/different or why they happened.
		0.8: High questioning , 36–45% of meeting codes include probing/reflective questions. High feedback , 36–45% of meeting codes include feedback with information to move practice forward.	0.8: High proximity to practice : 36–45% of activity code on discussion of potential uses of FA for student learning, teacher collaboration, or school-wide reform. (And/Or) 36–45% of depth of content discussed involving linking theory and practice.	0.8: High DoD : Almost all the way in the case of a team that has discussion time as spent involving discussion that: links ideas and examples through an examination of WHY things are similar/different or why they happened as an integral part of their meeting time. Graph indicates 10–20% of overall meeting time involving discussion that: links ideas and examples through an examination of WHY things are similar/different or why they happened.
	1: Single purpose	1: Very High questioning , 46% or more of meeting codes include probing/reflective questions. Very High feedback , 46% or more of meeting codes include feedback with information to move practice forward.	1: Very high proximity to practice : 46% or more of activity code on discussion of potential uses of FA for student learning, teacher collaboration, or school-wide reform. (And/Or) 46% or more of depth of content discussed involving linking theory and practice.	1: Very High DoD : Graph indicates that 21–100% of overall meeting time involving discussion that: links ideas and examples through an examination of WHY things are similar/different or why they happened. This case represents a high degree of discussion that probes deeper to whys as a regular part of their meeting. How we are defining regular is more than 21% of the time.

References

- Achen, C. H. (2005). Two cheers for Charles Ragin. *Studies in Comparative International Development (SCID)*, 40(1), 27–32.
- Antonacopoulou, E. (2004). On the virtues of practicing scholarship: a tribute to Chris Argyris, a 'Timeless Learner'. *Management Learning*, 35(4), 381–395.
- Avalos, B. (2011). Teacher professional development in teaching and teacher education over ten years. *Teaching and Teacher Education*, 27(1), 10–20.
- Bell, B., & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education*, 85, 536–553.
- Black, P., & Wiliam, D. (1998). *Inside the black box: Raising standards through classroom assessment*. Granada Learning.
- Borko, H. (2004). Professional development and teacher learning: mapping the terrain. *Educational Researcher*, 33(8), 3–15.
- Borko, H., & Putnam, R. (1995). Expanding a teachers' knowledge base: a cognitive psychological perspective on professional development. In T. Guskey, & M. Huberman (Eds.), *Professional development in education: New paradigms and practices* (pp. 35–66). New York: Teachers College Press.
- Boud, D., & Walker, D. (1998). Promoting reflection in professional courses: the challenge of context. *Studies in Higher Education*, 23(2), 191–206.
- Brockbank, A., McGill, I., & Beech, N. (2002). *Reflective learning in practice*. Gower Publishing, Ltd.
- Brookfield, S. (2005). *The power of critical theory. Liberating adult learning and teaching*. Jossey-Bass.
- Brookhart, S. M. (2013). Classroom assessment in the context of motivation theory and research. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom Assessment* (pp. 35–55). Thousand Oaks, CA: SAGE Publications, Inc.
- Bryk, A., Lee, V., & Holland, P. (1993). *Catholic schools and the common good*. Cambridge, MA: Harvard University Press.
- CCSSO. (2008). *Attributes of effective formative assessment*. Washington, DC: Council of Chief State School Officers.
- Cobb, P., McCain, 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.
- Coburn, C. E. (2001). Collective sense making about reading: how teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis*, 23(2), 145–170.
- Cochran-Smith, M., & Lytle, S. L. (1992). Communities for teacher research: fringe or forefront? *American Journal of Education*, 298–324.
- Cooper, B., & Glaesser, J. (2011). Paradoxes and pitfalls in using fuzzy set QCA: illustrations from a critical review of a study of educational inequality. *Sociological Research Online*, 16(3).
- Cooper, B. S., & Shear, R. G. (2012). *Screwed-up school reform: Fixing America's broken promise*. R&L Education.
- Costa, A. L., & Garmston, R. J. (1994). *Cognitive coaching: A foundation for renaissance schools*. 480 Washington, Street, Norwood, MA: Christopher-Gordon Publishers, Inc.
- Cuban, L. (2013). Why so many structural changes in schools and so little reform in teaching practice? *Journal of Educational Administration*, 51(2), 109–125.
- Curry, M. W. (2008). Critical friends groups: the possibilities and limitations embedded in teacher professional communities aimed at instructional improvement and school reform. *Teachers College Record*, 114(4), 733–774.
- Darling-Hammond, L. (1997). *The right to learn: A blueprint for creating schools that work*. San Francisco: Jossey-Bass.
- Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, N., & Orphanos, S. (2009). State of the profession: study measures status of professional development. *Journal of Staff Development*, 30(2), 42.
- Desimone, L. (2009). Improving impact studies of professional development: toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181–199.
- Desimone, L. M., Smith, T. M., Hayes, S., & Frisvold, D. (2005). Beyond accountability and average math scores: relating multiple state education policy attributes to changes in student achievement in procedural knowledge, conceptual understanding and problem solving in mathematics. *Educational Measurement: Issues and Practice*, 24(4), 5–18.
- Dobie, T. E., & Anderson, E. R. (2015). Interaction in teacher communities: three forms teachers use to express contrasting ideas in video clubs. *Teaching and Teacher Education*, 47, 230–240.
- Dooner, A. M., Mandzuk, D., & Clifton, R. A. (2008). Stages of collaboration and the realities of professional learning communities. *Teaching and Teacher Education*, 24(3), 564–574.
- DuFour, R., Eaker, R., & DuFour, R. (2005). *On common ground*. Bloomington: National Educational Service.
- DuFour, R., & Mattos, M. (2013). How do principals really improve schools? *Educational Leadership*, 70(7), 34–40.
- Education Week. (2014). *Diplomas count 2014*. Education Week Research Center. Retrieved from <http://www.edweek.org/ew/toc/2014/06/05/index.html>.
- van Es, E. A. (2012). Examining the development of a teacher learning community: the case of a video club. *Teaching and Teacher Education*, 28(2), 182–192.
- Fishman, J. J., Marx, R. W., Best, S., & Tal, R. T. (2003). Linking teacher and student learning to improve professional development in systemic reform. *Teaching and Teacher Education*, 19, 643–658.
- Fiss, P. C. (2007). A set-theoretic approach to organizational configurations. *Academy of Management Review*, 32(4), 1180–1198.
- Fiss, P. C. (2008). Institutions and corporate. In *The Sage handbook of organizational institutionalism*, 389.
- Franke, M. L., Carpenter, T. P., Feinman-Nemser, S., Ansell, E., & Behrend, J. (1998). Understanding teachers' self sustaining, generative change in the context of professional development. *Teaching and Teacher Education*, 14(1), 67–80.
- Frederiksen, J. R., Sipusic, M., Sherin, M., & Wolfe, E. W. (1998). Video portfolio assessment: creating a framework for viewing the functions of teaching. *Educational Assessment*, 5(4), 225–297.
- Furtak, E. M. (2012). Linking a learning progression for natural selection to teachers' enactment of formative assessment. *Journal of Research in Science Teaching*, 49(9), 1181–1210.
- Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38(4), 915–945.
- Grissom, J. A., & Herrington, C. D. (2012). Struggling for coherence and control: the new politics of intergovernmental relations in education. *Educational Policy*, 26(1), 3–14.
- Grossman, P., Wineburg, S., & Woolworth, S. (2001). Toward a theory of teacher community. *Teachers College Record*, 103(6), 942–1012.
- Hadar, L., & Brody, D. (2010). From isolation to symphonic harmony: building a professional development community among teacher educators. *Teaching and Teacher Education*, 26(8), 1641–1651.
- Horn, I. S., & Little, J. W. (2009). Attending to problems of practice: routines and resources for professional learning in teachers' workplace interactions. *American Educational Research Journal*, 47(1).
- Katz, S., Sutherland, S., & Earl, L. (2005). Toward an evaluation habit of mind: mapping the journey. *The Teachers College Record*, 107(10), 2326–2350.
- King, M. B. (2002). Professional development to promote school wide inquiry. *Teaching and Teacher Education*, 18(3), 243–257.
- King, M. B., & Newmann, F. M. (2001). Building school capacity through professional development: conceptual and empirical considerations. *International Journal of Educational Management*, 15(2), 86–94.
- Labaree, D. F. (2003). The peculiar problems of preparing educational researchers. *Educational Researcher*, 32(4), 13–22.
- Lachance, A., & Confrey, J. (2003). Interconnecting content and community: a qualitative study of secondary mathematics teachers. *Journal of Mathematics Teacher Education*, 6, 107–137.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.
- Levine, T. H. (2010). Tools for the study and design of collaborative teacher learning: the affordances of different conceptions of teacher community and activity theory. *Teacher Education Quarterly*, 109–130.
- Lieberman, A. (1995). *Practices that support teacher development: Transforming conceptions of professional learning* (p. 67). Innovating and Evaluating Science Education. NSF Evaluation Forums, 1992–94.
- Lieberman, A., & Mace, D. H. P. (2009). The role of 'accomplished teachers' in professional learning communities: uncovering practice and enabling leadership. *Teachers and Teaching: Theory and practice*, 15(4), 459–470.
- Lord, B. (1994). Teachers' professional development: critical collegueship and the role of professional communities. *The Future of Education: Perspectives on National Standards in Education*, 175–204.
- Lotter, C., Harwood, W. S., & Bonner, J. J. (2006). Overcoming a learning bottleneck: inquiry professional development for secondary science teachers. *Journal of Science Teacher Education*, 17(3), 185–216.
- Loucks-Horsey, S., Hewson, P. W., Love, N., & Stiles, K. E. (1998). *Designing professional development for teachers of science and mathematics*. Thousand Oaks, CA: Corwin Press.
- Louis, K. S., & Marks, H. M. (1998). Does professional community affect the classroom? Teachers' work and student experiences in restructuring schools. *American Journal of Education*, 532–575.
- Marshall, B., & Drummond, M. J. (2006). How teachers engaged with assessment for learning: lessons from the classroom. *Research Papers in Education*, 21(2), 133–149.
- McIntyre, E., & Kyle, D. W. (2006). The success and failure of one mandated reform for young children. *Teaching and Teacher Education*, 22(8), 1130–1144.
- McLaughlin, M. W. (1991). The Rand change agent study: ten years later. *Education Policy Implementation*, 143–155.
- McLaughlin, M. W. (2005). Listening and learning from the field: tales of policy implementation and situated practice. In *The roots of educational change* (pp. 58–72). Netherlands: Springer.
- Michigan Department of Education. (2011). *Formative assessment research report*. Retrieved at http://www.michigan.gov/mde/0,4615,7-140-22709_55936--,00.html.
- Miles, M. B., & Huberman, A. M. (1994). *An expanded sourcebook: Qualitative data analysis* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards*. Washington, DC: Authors.
- Nelson, T., Slavit, D., Perkins, M., & Hathorn, T. (2008). A culture of collaborative inquiry: learning to develop and support professional learning communities. *The Teachers College Record*, 110(6), 1269–1303.
- Newmann, F., King, M. B., & Youngs, P. (2000). Professional development that addresses school capacity: lessons from urban elementary schools. *American Journal of Education*, 108(4), 259–299.
- Newmann, F. M., Smith, B., Allensworth, E., & Bryk, A. S. (2001). Instructional program coherence: what it is and why it should guide school improvement policy.

- Educational Evaluation and Policy Analysis, 23(4), 297–321.
- Penuel, W. R., Fishman, B. J., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional strategies development effective? That Foster curriculum implementation. *American Educational Research Journal*, 44(4), 921–958.
- Popham, W. J. (2009). Assessment literacy for teachers: faddish or fundamental? *Theory into Practice*, 48(1), 4–11.
- Ragin, C. C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies*. Berkeley: University of California Press.
- Ragin, C. C. (2000). *Fuzzy-set social science*. Chicago: University of Chicago Press.
- Ragin, C. C. (2008). *Redesigning social inquiry: Fuzzy sets and beyond* (pp. 190–212). Chicago: University of Chicago Press.
- Ragin, C. C. (2014). *The comparative method: Moving beyond qualitative and quantitative strategies*. Berkeley: University of California Press.
- Ragin, C. C., & Fiss, P. C. (2008). Net effects analysis versus configurational analysis: An empirical demonstration. In *Redesigning social inquiry: Fuzzy sets and beyond* (pp. 190–212).
- Richmond, G., & Manokore, V. (2011). Identifying elements critical for functional and sustainable professional learning communities. *Science Education*, 95, 543–570.
- Rihoux, B., & Lobe, B. (2009). The case for qualitative comparative analysis (QCA): adding leverage for thick cross-case comparison. In D. Byrne, & C. C. Ragin (Eds.), *The Sage handbook of case-based methods* (pp. 222–242). Sage Publications Ltd.
- Schneider, M. C., & Randel, B. (2009). Research on characteristics of effective professional development programs for enhancing educators' skills in formative assessment. In H. L. Andrade, & G. Cizek (Eds.), *Handbook of formative assessment* (pp. 251–276). New York, NY: Routledge.
- Sebastian, J., Allensworth, E., & Stevens, D. (2014). The influence of school leadership on classroom participation: examining configurations of organizational supports. *Teachers College Record*, 116(8), 1–36.
- Skerrett, A. (2010). "There's going to be community. There's going to be knowledge": designs for learning in a standardised age. *Teaching and Teacher Education*, 26(3), 648–655.
- Slavit, D., Nelson, T. H., & Kennedy, A. (2010). Laser focus on content strengthens teacher teams. *Journal of staff development*, 31(5), 18.
- Slavit, D., Laurence, W., Kennedy, A., & Nelson, T. H. (2009). 7 resource networks for collaborative teacher inquiry. In D. Slavit, T. Holmlund Nelson, & A. Kennedy (Eds.), *Perspectives on supported collaborative teacher inquiry* (p. 145). New York: Taylor & Francis.
- Snow-Geron, J. L. (2005). Professional development in a culture of inquiry: PDS teachers identify the benefits of professional learning communities. *Teaching and Teacher Education*, 21(3), 241–256.
- Stake, R. E. (2013). *Multiple case study analysis*. Guilford Press.
- Stein, M. K., Smith, M. S., & Silver, E. A. (1999). The development of professional developers: learning to assist teachers in new settings in new ways. *Harvard Educational Review*, 69(3), 237–269.
- Stoll, L., Bolam, R., McMahon, A., Wallace, M., & Thomas, S. (2006). Professional learning communities: a review of the literature. *Journal of Educational Change*, 7, 221–258.
- Strauss, A., & Corbin, J. (1990). *Basics of qualitative Research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage Publications.
- Supovitz, J. (2002). Developing communities of instructional practice. *The Teachers College Record*, 104(8), 1591–1626.
- Talbert, J., Wallin McLaughlin, M., & Rowan, B. (1993). Understanding context effects on secondary school teaching. *The Teachers College Record*, 95(1), 45–68.
- Taylor, K. L., & Colet, N. R. (2010). Making the shift from faculty development to educational development. *Building Teaching Capacities in Higher Education*, 139.
- Thomas, G., Wineburg, S., Grossman, P., Myhre, O., & Woolworth, S. (1998). In the company of colleagues: an interim report on the development of a community of teacher learners. *Teaching and Teacher Education*, 14(1), 21–32.
- Trujillo, T. M., & Woulfin, S. L. (2014). Equity-oriented reform amid standards-based accountability: a qualitative comparative analysis of an intermediary's instructional practices. *American Educational Research Journal*, 51(2), 253–293.
- Vescio, V., Ross, D., & Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and Teacher Education*, 24(1), 80–91.
- 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* (p. 129). Dallas, TX.
- Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge, UK: Cambridge University Press.
- William, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, 37(1), 3–14.
- Wilson, S. M. (2013). Professional development for science teachers. *Science*, 340, 310–313.
- Wilson, S. M., & Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: an examination of research on contemporary professional development. *Review of Research in Education*, 24, 173–209.
- Wylie, C., Lyon, C., & Goe, L. (2009). *Teacher professional development focused on formative assessment: Changing teachers, changing schools*. Research Report. Princeton, NJ: Educational Testing Service.
- Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). Washington D.C.: Sage.
- Yoon, K. S., Duncan, T., Lee, S. W. Y., Scarloss, B., & Shapley, K. L. (2007). *Reviewing the evidence on how teacher professional development affects student achievement*. Issues & Answers. REL 2007-No. 033. Southwest (NJ1): Regional Educational Laboratory.
- Youngs, P., & King, M. B. (2002). Principal leadership for professional development to build school capacity. *Educational Administration Quarterly*, 38(5), 643–670.