Student teachers' and in-service teachers' peer learning

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.
Link to publication

Student teachers' and in-service teachers' peer learning: a realist synthesis

Marieke Thurlings & Perry den Brok

Published online: 19 Aug 2018.

Submit your article to this journal ⬈

Article views: 458

View Crossmark data ⬈

Routledge
Taylor & Francis Group

Student teachers' and in-service teachers' peer learning: a realist synthesis

Marieke Thurlings[a] and Perry den Brok[b]

[a]Eindhoven School of Education, Eindhoven University of Technology, Eindhoven, The Netherlands; [b]Education and Learning Sciences, Wageningen University & Research, Wageningen, The Netherlands

ABSTRACT

The aim of this review in the form of a realist synthesis is to understand what makes peer (student) teacher learning activities effective. Three types of peer learning were explored: collaborative, coaching, and assessment activities. Hypotheses were developed and tested against a set of 63 studies. These hypotheses indicated what mechanisms (i.e., characteristics of the activities) would influence the effectiveness of the activities and which contexts (i.e., factors) influenced this process. Findings showed that activities wherein participants gain subject matter and pedagogical content knowledge and apply this into practice made such learning effective. Peer learning is also reinforced by facilitators, reflection, and feedback, and influenced by personal and interpersonal factors. The main contribution of this realist synthesis is the practical implications for developers of peer learning activities, for school leaders and teacher educators, and for (student) teachers.

Introduction

It is widely acknowledged that teacher professional development is essential in today's educational system. Teachers continuously need to update their knowledge and instruction skills in order to meet the demands of society. For example, they need to be able to use technology in their classrooms and adapt to changing student populations (Thurlings, Evers, & Vermeulen, 2015). Previous literature reviews on teacher learning (e.g., Desimone, 2009; Van Veen, Zwart, & Meirink, 2012) have shown that learning from and with peers is an effective approach to teacher professional development. Here, we refer to these kind of activities as peer teacher learning (PTL) activities, where peer means of equal status. Van Veen et al. (2012) recommended, on the basis of findings of their literature review about teacher professional development interventions, that more research is needed on what characteristics make these activities effective and which school organisation conditions are needed for effective PTL.

Moreover, more and more inititiatives are taken wherein teachers work together (Schleicher, 2016); yet, these are not always optimal in reaching their goals, due to, for

example, constraints in the organisation or personal issues. Insights into organisational and personal factors that influence the effectivity of PTL activities and the characteristics of those activities can support schools, school leaders, and teacher education institutes to improve these issues. To provide such a basis of evidence for both research and practice, the present review combines a systematic literature review with a focus on formulating such factors and characteristics, which can be used for developing and evaluating teacher professional development.

The present study provides a review of literature on PTL in the form of a realist synthesis (Pawson, Greenhalgh, Harvey, & Walshe, 2005), which is used more and more, especially in the medical domain (Rycroft-Malone et al., 2012). Realist syntheses are grounded in realism as a philosophy to science (Dalkin, 2014; Dalkin, Greenhalgh, Jones, Cunningham, & Lhussier, 2015; Wong, Greenhalgh, Westhorp, Buckingham, & Pawson, 2013; Wong, Westhorp, Pawson, & Greenhalgh, 2013). Realism is positioned between positivism and constructivism, and agrees that a real world exists and that humans construct knowledge within this real world. However, realism postulates that a better understanding of reality is possible, because the "real world" hinders interpretation. In order to reach this better understanding of reality, realist synthesis begins by formulating programme theories that reason why things work (Dalkin, 2014), delving into the black box of input and output. As such, programme theories describe mechanisms (i.e., characteristics of and resources within the PTL activity). Additionally, these mechanisms within the PTL activity influence the behaviour of participants, adding a second layer to these mechanisms (Dalkin et al., 2015). The context of an intervention then influences these layers of mechanisms, leading to outcomes. These context-mechanism-outcome configurations are tested against included studies and are refined throughout this process (Dalkin, 2014). As such, realist synthesis is a "policy-friendly" (Wong, Greenhalgh, et al., 2013, p. 11) approach to literature reviews.

The aim of the present realist synthesis is to analyse and synthesise what mechanisms and contexts lead to positive learning outcomes, such as knowledge and skills, in a variety of PTL activities. The second aim is to formulate practical implications (in other words, context-mechanism-outcome configurations) that can be used to developed and evaluate PTL activities. The scope of the realist synthesis is both on student teachers and in-service teachers in primary and secondary education. As such, the realist synthesis approach enables us to explore which specific characteristics of PTL activities make them effective, and to provide directions for evaluating PTL activities.

Realist synthesis, as an approach to synthesising literature, was originally developed by Pawson (2002) in the health care domain. Pawson (2002; Pawson et al., 2005) suggested that a realist synthesis can also be applied to other fields, such as education, because interventions have similar characteristics as in health care. Similar to health care, PTL interventions are situated in a larger, societal context. While implementing PTL interventions, contexts and mechanisms such as the school leader, daily practice, and personal factors of involved teachers are likely to influence the success of the PTL activity. Realist synthesis caters to this complexity. Moreover, realist synthesis supports the formulation of practical implications, enabling to bridge the gap between research and practice.

There are many approaches to and types of PTL activities. These activities have many different names and definitions, some of which are ill defined and sometimes overlapping (see also Vangrieken, Dochy, Raes, & Kyndt, 2015). A well-known conceptualisation is that

of Little (1990), who distinguished a continuum of interdependency in interactions: (a) storytelling and scanning for ideas, where teachers informally share stories; (b) aid and assistance, where teachers support each other when asked for; (c) sharing, where teachers, for example, provide materials to each other in a one-way street; and (d) joint work, where teachers collaborate and build on each other's expertise. In the present literature review, on the basis of functions that activities may serve as well as the size of the group involved in that activity, this variety in PTL activities is conceptually divided into three main types: coaching, collaborating, and assessing PTL activities.

Coaching PTL activities are defined here as activities wherein individuals support another individual, reciprocal or non-reciprocal, in achieving individual goals, for example, when teachers provide peer feedback to each other. They might peer coach or peer mentor each other or participate in collegial consultation. This PTL activity can be compared to Little's (1990) category of aid and assistance. Previous literature studies focused on similarities and differences between peer coaching programmes (Ackland, 1991; Lu, 2010). Ackland (1991) found three common characteristics of these programmes. The first characteristic is that peer coaching is non-evaluative: that is, no decisions are made about the performance of the participating teachers, and, as a consequence, peer coaching is a safe environment. Second, peer coaching is based on classroom observations followed by constructive feedback. In some peer coaching programmes, these observations involve videotaping the coached teachers' classrooms. Third, effective peer coaching is directed at improving instructional techniques and teacher behaviours.

Lu (2010) focused on peer coaching in teacher education and revealed similarities and differences between eight peer coaching programmes for student teachers. Concerning similarities, Lu showed (a) participants were often only student teachers; (b) only peers were involved in coaching; (c) peer coaching aligned with student teachers' field experiences; and (d) peer coaching often encompassed pre-observation sessions, observations, and post-observations sessions. Concerning differences, Lu found that (a) the focus of peer coaching ranged from professional development to field experiences to still other aspects; (b) the context differed, for example, from peer learning by special education teachers to physical education teachers; (c) the training student teachers received differed in duration and content; and (d) effects differed, from greater perceived importance of field experiences to a higher perceived amount of affective support of peer coaches.

The second type of PTL activities this literature review focuses on are *collaborative activities*, defined in line with Vangrieken et al. (2015) as activities wherein larger groups (than in coaching activities) collaborate on a shared goal and/or product. In coaching groups, there are two to four members; in collaborative groups, four or more (up to nine in the included studies). For example, teachers might collaborate in teams or in networks, or they might participate in lesson study teams or in a teacher design team. This PTL activity can be compared to Little's (1990) category of joint work. Vescio, Ross, and Adams (2008) explored the impact of teacher communities of practice on teacher practices and student learning in 11 studies. Five of these studies indicated changes in teacher practices, and eight studies revealed changes in student learning. All studies reported on changes in the school's professional culture, and Vescio et al. showed that several characteristics promoted these changes: Collaboration encouraged "sharing, reflecting and taking risks necessary to change" (p. 84); the focus of collaboration was on student learning;

participating teachers were able to make decisions about processes in their communities and certain aspects of the school organisation; and community participation supported continuous learning by teachers. A recent review focused on teacher collaboration (Vangrieken et al., 2015) indicated a variety of positive outcomes of teacher collaboration, such as student learning and improved teacher practice (cf. Vescio et al., 2008). Additionally, Vangrieken et al. revealed a range of facilitating and hindering factors, such as personal characteristics of teachers and organisational characteristics.

The third type of PTL activity this literature review focuses on is *assessment activities*, which refers to activities wherein individuals assess another individual's performance in either formative or summative ways. Teachers might participate in peer reviewing or peer assessment. There are several literature reviews on peer assessment (e.g., Dochy, Segers, & Sluijsmans, 1999; Topping, 1998; Van Gennip, Segers, & Tillema, 2009), but neither of these focused exclusively on teachers or student teachers. Dochy et al. (1999) reviewed self- and peer assessment procedures and showed that peer assessment has several benefits, such as higher student involvement in learning processes. Topping (1998) provided a clear definition and an overview of characteristics of peer assessment. Additionally, he showed that a variety of factors influence the application of peer assessment, such as clarifying expectations, clear criteria, and skill training. Van Gennip et al. (2009) intended to explore how interpersonal factors, such as trust, influence peer assessment, but found that only a few studies investigated this relationship.

To summarise the three PTL activities: In collaborative activities, (student) teachers work together to reach a shared goal; in coaching activities, they support each other's process in reaching each individual's goal; and in assessment activities, they judge the others' performance.

None of the literature studies described above focused on specific characteristics of PTL activities that make these activities work. Another issue with former literature reviews is that they were limited by focusing on only one type of PTL activities. A third issue is that these literature studies were limited to in-service teachers (Ackland, 1991; Vangrieken et al., 2015; Vescio et al., 2008) or to student teachers (Lu, 2010), and in the case of peer assessment, they did focus on neither in-service nor or student teachers. In short, former literature studies were restricted because they focused on just one PTL activity, looked at either pre-service or in-service teachers, and provided no clear insights into which characteristics of the PTL activity made it effective. The effectiveness of PTL activities is seen as whether mechanisms and contexts contribute to performing, executing, and applying that activity (Dalkin, 2014; Darling-Hammond & Snyder, 2000). To address the issues with former literature reviews as well as the practical needs, the realist synthesis approach is suitable.

The main research question was "Under which circumstances do PTL activities work for (student) teachers and why?" This main research question can be divided into four sub-questions. The first two focus on the mechanisms of PTL activities and the contexts that influence the effectiveness. The second and third questions focus on differences between circumstances. The fourth question focuses on differences between teachers and student teachers:

(1) Which mechanisms of PTL activities enhance or hinder the effectiveness of peer (student) teacher learning (PTL) activities?

(2) Which contexts enhance or hinder the effectiveness of peer (student) teacher learning (PTL) activities?
(3) Are there differences in enhancing and hindering characteristics and factors for different types of PTL, namely, collaborative, coaching, and assessment activities?
(4) Are there differences in enhancing and hindering characteristics and factors between peer learning by student teachers and by in-service teachers?

Elsewhere, we have reported on the outcomes of PTL activities and the strengths and weaknesses of research focusing on these outcomes (Thurlings & Den Brok, 2017).

Methods

In this section, we first explain the procedures for conducting a realist synthesis because these procedures differ somewhat from regular literature review approaches. Next, we describe how we conducted the realist synthesis in the present review.

The general procedures of realist synthesis

Realist synthesis is an approach to synthesising both quantitative and qualitative studies (Pawson, Greenhalgh, Harvey, & Walshe, 2004). The aim of realist synthesis is to answer the question: What works for whom under what circumstances? Generally, realist synthesis applies the same steps that a conventional literature review would; however, it has some distinctive features. A conventional literature review (e.g., Petticrew & Roberts, 2006) starts by formulating research questions. Subsequently, conventional literature reviewers search and select studies, and next, they extract data and appraise the quality of studies. The next step is to synthesise the findings, and the final step is to disseminate findings.

A realist synthesis also begins by formulating a research question (Pawson et al., 2004, 2005). Second, a programme theory (in our case a set of hypotheses) is developed that explains why an intervention could be effective. These hypotheses describe which mechanisms, together with which contextual factors, influence the effectiveness of the interventions (Dalkin, 2014; Pawson et al., 2004, 2005). The section Developing and testing hypotheses describes how we approached this step. The third step is to search and select studies (see section Searching and selecting studies). Next, data are extracted from the included studies that are used to test the hypotheses that were developed in the second step. Realist synthesis does not necessarily appraise the quality of studies (Pawson et al., 2004), yet it asks if the studies are relevant to answer the research questions and whether the evidence has enough rigour to draw conclusions. Relevance means that the included studies address the hypotheses that are being tested; rigour means that findings are credible and providing sufficient weight in terms of method/methodology (Pawson et al., 2004). The fifth step is to actually test the hypotheses with the extracted evidence, which leads to refining the context-mechanisms-outcomes configurations. These refined configurations can be considered as practical recommendations. As a sixth step, Pawson et al. (2004, 2005) recommend discussing the context-mechanism-outcome configurations with actors and stakeholders. An advantage of such a discussion is that the hidden mechanisms (the behaviour of participants within the intervention) can

be further elicited (Dalkin, 2014) and that it also builds a bridge between research and practice. The section Stakeholders describes how we discussed our findings with stakeholders. Finally, the seventh step contains the actual dissemination of findings. In writing this article, we applied the RAMESES publication standards (Wong, Greenhalgh, et al, 2013).

Developing hypotheses

The hypotheses, or programme theory, we developed are based on former literature studies that have shed light on why and how professional development and PTL activities could work for (student) teachers' learning. These literature studies (e.g., Desimone, 2009; Van Veen et al., 2012) were read and their findings and conclusions summarised. Similar findings and conclusions were compiled within an inventory. For example, both Desimone (2009) and Van Veen et al. (2012) concluded that PTL activities should focus on gaining subject matter and/or pedagogical content knowledge. Thus, this exploration of former literature gave us a preliminary inventory about why and how PTL activities could be effective.

The inventory was labelled into mechanisms and contexts. Mechanisms were seen as characteristics of the processes (e.g., effective PTL activities focus on gaining subject matter and/or pedagogical content knowledge, based on Capps, Crawford, and Constas, 2012; Desimone, 2009; and Van Veen et al., 2012), and contexts were considered as factors that influence the processes. Finally, 10 hypotheses were formulated (see Table 1); for instance, PTL activities become more effective when they show a gain in subject matter and/or pedagogical content knowledge (#1) than if they do not show a gain in such knowledge. The hypotheses expressed why and how PTL activities could be effective for (student) teachers' learning: if they included given aspects or tools, they would be more effective (Research question 1). Hypotheses 1 through 7 describe mechanisms, while Hypotheses 8, 9, and 10 describe contexts. The hypotheses concerning factors (i.e., 8, 9, and 10) were neutrally formulated; that is, it was hypothesised that factors would influence PTL activities (Research question 2). During analyses, it was explored whether this influence was positive (promoting PTL activities) or negative (hindering PTL activities), and in the Results section, these findings are exemplified.

Searching and selecting studies

The search was conducted throughout the summer and autumn of 2013, right after the grant (see Funding) was awarded. Late autumn and during the winter of 2013–2014, studies were further selected. During the spring and summer of 2014, data were extracted and analysed. At the end of summer 2014, a symposium with stakeholders was organised, after which the manuscript was written. Studies published between January 1991 and the summer and autumn of 2013 were explored. To locate relevant studies, three search strategies were used. The first strategy was to explore two search engines, namely, EBSCO host and Science Direct. A variety of search terms was used, such as *collaborative learning*, *Community of Practice*, *peer coaching*, *peer feedback*, *peer assessment*, and *peer review*. As alternatives for *peer*, we also used *colleague*, *equal*, and *collegial* as search terms. Each

Table 1. The hypotheses to be tested in the realist synthesis.

#	Hypothesis	Number of data-sets that addressed the hypothesis	Literature used to phrase the hypothesis
	Mechanisms of the PTL activities		
1	PTL activities become more effective when they show a gain in subject matter and/or pedagogical content knowledge	27	Capps et al. (2012). Desimone (2009), Van Veen et al. (2012)
2	PTL activities become more effective when they encompass active learning activities, such as observing or analysing, rather than passive activities, such as listening	17	Desimone (2009), Lauer, Christopher, Firpo-Triplett, and Buchting (2014), Van Veen et al. (2012)
3	PTL activities become more effective when they have a reflection component	18	Ackland (1991), Capps et al. (2012)
4	PTL activities become more effective when they have a feedback component	9	Ackland (1991)
5	PTL activities become more effective when they apply video fragments of teaching	9	Borko, Koellner, and Jacobs (2011), Brophy (2003)
6	PTL activities become more effective when they are guided by a facilitator or group leader	17	Ackland (1991), Borko et al. (2011), Lauer et al. (2014), Van Veen et al. (2012)
7	PTL activities become more effective when they have a longer rather than shorter duration and follow-ups.	13	Desimone (2009), Lauer et al. (2014), Van Veen et al. (2012)
	Contexts influencing PTL activities		
8	PTL activities become more effective when personal factors, such as motivation and beliefs, are taken into account.	39	Desimone (2009), Evers, Kreijns, Van der Heijden, & Gerrichhauzen (2011)
9	PTL activities become more effective when interpersonal factors, such as trust and respect, are taken into account.	36	Capps et al. (2012), Salas, Sims, and Burke (2005), Van Gennip et al. (2009)
10	PTL activities become more effective when contextual factors, such as support from school leaders and ICT, are taken into account.	22	Avalos (2011), Van Veen et al. (2012)

search term was additionally combined with either *teacher*, *student teacher*, *teacher education*, or *professional development*.

Second, we manually searched 11 journals that specifically focus on (student) teachers' learning and development and were therefore expected to contain relevant studies. These journals were, in alphabetical order: *Action in Teacher Education*, *Asia-Pacific Journal of Teacher Education*, *European Journal of Teacher Education*, *Journal of Education for Teaching*, *Journal of Teacher Education*, *Teacher Development*, *Teacher Education Quarterly*, *Teachers and Teaching*, *Teachers College Record*, *Teaching and Teacher Education*, and *Professional Development in Education*. Third, we searched PhD dissertations. Additional criteria for the inclusion of PhD dissertations were that these should have been assessed by an assessment committee and should be available online.

To focus the search, we decided upon a number of criteria for inclusion. These were as follows: (a) The article reported on an empirical study; (b) the participants were teachers or student teachers in initial education (i.e., primary, secondary, and special education); (c) participants were peers; (d) learning was explored in terms of processes and activities in a formal setting; (e) the study's findings were helpful in answering at least one of our research questions; and (f) the study was not included in the former literature studies used to develop the 10 hypotheses. Studies were excluded if they did not meet one, or more, of these criteria for inclusion.

Concerning criterion (c), peers were defined as being of the same or similar status (cf. Ackland, 1991). More specifically, this meant that student teachers should have the same amount of field experience and in-service teachers should have the same position in the PTL process. We exemplify this delineation with one of the included studies. Vincent and Jones (2008) examined a coaching PTL activity at a school that had provided each classroom with an interactive whiteboard. Two teachers were relieved of teaching tasks in order to coach other teachers in using the interactive whiteboard. These two teachers did, like their colleagues, have no experience whatsoever in using interactive whiteboards. Therefore, they were peers according to our definition. If the two had had experience in teaching with interactive whiteboards, they would not have been peers, and the study would have been excluded. If it was too unclear whether the (student) teachers in a study were peers according to our definition, the study would have been excluded. Concerning criterion (e), findings of the included studies had to be relevant, which means that the findings from the included studies had to address the hypotheses tested (Pawson et al., 2004).

We did include articles in which a group of (student) teachers (i.e., peers) was guided by a teacher educator or researcher. However, such guidance had to take the form of moderating or making sure the peer learning process was fluent, rather than focusing on the content of the activity itself. We chose to focus on peers only because the literature on PTL activities is vast and we needed a clear focus. In addition, the operationalisations of peers were chosen such that the participants would be colleagues in the simplest meaning of the word.

The selection process consisted of two steps. First, approximately 65,000 titles and abstracts (including duplicates) were judged against the six criteria and selected for further reading if they met at least five out of six criteria. About 950 publications were selected for further reading, where whole publications were again judged against the criteria and, in order to be selected, had to meet each single criterion. Ultimately, 69 publications were included.

Data extraction and hypotheses testing

After studies were searched and selected, evidence was extracted from the included studies to support the hypotheses. As "Pawson's publications provide little guidance on how to approach data synthesis" (Rycroft-Marlone et al., 2012, p. 6), we developed our own strategy. Each publication was carefully read, and we searched for text in the results and conclusion and discussion sections that connected to the 10 hypotheses. This text was copy pasted into Excel sheets, and for each hypothesis we used a single Excel sheet. To this text, we added a column that showed the type(s) of data collection the finding was revealed with (e.g., questionnaire, interview, or observation). In alignment with the realist synthesis approach, we also looked for contradicting evidence for the hypotheses (Pawson et al., 2004). In some cases, findings from included studies added to the 10 hypotheses, and these were also included into the Excel sheets.

After all evidence for all hypotheses was collected, to answer the first research question findings from Hypotheses 1 through 7 were interpreted to examine whether the hypotheses were supported or rejected. To answer the second research question, findings from Hypotheses 8 through 10 were interpreted, and categories of factors were established bottom-up.

To explore the third and fourth research questions, colour coding was used. Six different colour codes were created (collaboration activities & in-service teachers, coaching activities & in-service teachers, assessment activities & in-service teachers, collaboration activities & student teachers, coaching activities & student teachers, assessment activities & student teachers). As such, we could examine whether different hypotheses held true for each of these six settings. Finally, during the interpretation process, we took into account how findings were originally revealed by examining which type of instrument was used (e.g., survey, interview), and whether it was quantitative or qualitative in nature. As such, the rigour of findings was also taken into account (Pawson et al., 2004).

During the process of data extraction, testing the hypotheses, and answering the four research questions, we worked in close collaboration. The second author scored three of the 63 included studies into the 10 hypotheses, and as there was full agreement, the remaining studies were all scored by the first author. During the interpretation of findings, we held several meetings to discuss interpretations, aiming at making better informed decisions. In general, the first author conducted each step of the review and wrote the article. The second author had a supervising role throughout the whole review, including providing detailed and elaborate feedback and comments on the article. The second author also verified the coding process. The authors met regularly to discuss issues and make decisions.

Overview of studies

During reading the publications and extracting data, we found some publications that were based on the same or were part of one data collection. For instance, Meirink (2007; Meirink, Imants, Meijer, & Verloop, 2010; Meirink, Meijer, & Verloop, 2007; Meirink, Meijer, Verloop, & Bergen, 2009) investigated the same teachers; however, each publication focused on different research questions and on a different set of measurements. These kinds of publications were therefore considered as one record. As a consequence, 63 different data-sets were left for further analysis. In the remaining section, we provide an overview of these 63 data-sets; in the Results section, we refer to the individual publications wherein specific findings were revealed. Appendix 1 provides detailed information about the 63 data-sets: the country the study was undertaken in, the number and type of participants, the type of PTL activity, and the nature of the whole study (quantitative, qualitative, or mixed methods). Additionally, Appendix 1 shows per study to which of the hypotheses it linked and through which types of instruments these results were revealed.

Most studies were published in journals on teacher education or teacher development, and the journal *Teaching and Teacher Education* was represented the most. Also, several studies were published in ICT-related journals. Other journals were domain specific (e.g., *Journal of Literacy Research*) or focused on higher education (e.g., *Assessment & Evaluation in Higher Education*).

The studies were conducted throughout the world. The United States was represented most: 25 studies were situated therein. Across Europe, 18 studies were conducted, and eight studies were situated in Asia. Nine studies were conducted in Australia and one in New Zealand. One study was situated in Africa, and from one study it was unclear what its origin was.

Studies varied in their methodology. Some explored only one group of (student) teachers, others focused on several groups. All studies were based on interventions, and the majority examined participants in this intervention. Only a few studies compared two different interventions, and a few compared an intervention group (a new kind of teaching and learning) to a regular group (the usual kind of teaching and learning). Data were collected prior to, during, and after interventions, and all combinations of these were present. The majority of studies used a range of self-report data, such as questionnaires, interviews, and logbooks. More objective measures, such as observations, were used to a much lesser extent (see Appendix 1). Additionally, 43 studies were of a qualitative nature, 16 studies were characterised as a mixed method design, and four studies were of a quantitative nature.

The overall quality of studies was questionable. Especially in qualitative studies, the method sections did not provide much information, especially on how data were collected and analysed. Moreover, many studies relied only or mostly on self-reported data.

Forty-one studies focused on student teachers, 21 on in-service teachers, and in Grion and Varisco's study (2007), both student teachers and in-service teachers participated, yet in different groups. More specifically, 25 studies focused on collaboration activities with student teachers and 15 studies on collaboration activities with in-service teachers, and in one study both student and in-service teachers participated in collaboration activities (Grion & Varisco, 2007). Fourteen studies focused on coaching activities with student teachers, six on coaching activities with in-service teachers, and two studies focused on peer assessment activities with student teachers.

We did not find any relevant study on peer teacher assessment activities, which limits our realist synthesis. The literature on (student) teacher assessment activities is vast, yet the overwhelming majority of these studies explored participants' satisfaction with the assessment activities and/or investigated to what extent the judgement given was similar for peer, self-, and supervisor assessment. These kinds of questions are not relevant to address our research questions. Because we did find two relevant studies in three publications (Buchanan, 2011; Buchanan & Stern, 2012; Sluijsmans, Brand-Gruwel, Van Merriënboer, & Martens, 2004) on peer student teacher assessment activities, we decided to include these studies.

Stakeholders

Following recommendations of Pawson et al. (2004, 2005) and Dalkin (2014), we discussed the findings of the realist synthesis with stakeholders. This symposium took place at the end of September 2014. The symposium aimed to present the findings of the realist synthesis to representatives of schools (e.g., teachers and school management), teacher education institutes (e.g., teacher educators), and other organisations in the field of education, such as policymakers and educational advisors. The first author gave a presentation, outlining the goals, method, and major results. Two discussants were involved. One was a representative of secondary schools, had been a teacher, and was involved in management tasks and professional development projects. The second discussant was an assistant professor affiliated with another university than that of the authors, and a specialist in research on collaborative teacher learning. The discussants were provided with a preview

of the results. After their contributions, a discussion with the attendees of the symposium was held. This discussion centred on the findings that were just presented by the first author and discussed by the two discussants.

Approximately 40 people attended the symposium, and their participation was completely voluntary. They were affiliated to different organisations, such as universities, universities of applied sciences/colleges, schools, school boards, and non-profit interest groups. Their functions varied as well: policymakers, teachers, teacher educators, researchers, and school leaders. Some attendees combined functions; for example, they were teacher educator and researcher. During the discussion, they were free to contribute, and consequently, some did and others did not contribute to the discussion.

The findings were recognised by the discussants as well as by the audience. The second discussant suggested incorporating type of school (primary or secondary education) into the analysis. We explored this suggestion; however, we had to conclude that this information was not included in all studies and therefore we were not able to incorporate this suggestion.

Results

This section reveals the results of the testing of the 10 hypotheses as formulated in Table 1. The two subsections focus respectively on the mechanisms of PTL activities (first research question) and on the contexts influencing PTL (second research question). The third and fourth research questions are interwoven with the hypotheses and are discussed within each subsection. We first describe the main results per hypothesis, as found in the majority of studies, and provide illustrative examples of these main results. Second, we discuss exceptions to the main results and focus on, where applicable, differences found between types of activities (i.e., collaborative, coaching, and assessment activities) and between participants (i.e., student teachers and in-service teachers). In the case of additional findings (adding to the original hypothesis), these findings are addressed as such and described.

Mechanisms of PTL activities

Hypothesis 1: PTL activities become more effective when they show a gain in subject matter and/or pedagogical content knowledge

Regardless of the type of PTL and type of participants, all studies linked to this hypothesis focused on gaining subject matter and/or pedagogical content knowledge as an outcome of PTL. For instance, teachers reported that collaboration in teams helped them prepare and use new materials, which helped them applying new knowledge (Estebaranz, Mingorance, & Marcelo, 1999). Primary school teachers reported after participating in lesson study – where groups of teachers work together to design, implement, evaluate, and adapt lessons – that they had gained knowledge about, for example, instructional techniques (i.e., pedagogical content knowledge; Rock & Wilson, 2005). Student teachers reported after participating in collaborative groups that they had gained, for example, understanding of what it means to be a teacher (Eteläpelto, Littleton, Lahti, & Wirtanen, 2005). In another study where primary school teachers collaborated on implementing assessment practices, constructing new ideas and practices

were considered outcomes of teacher group discussions (Borko, Mayfield, Marion, Flexer, & Cumbo, 1997; see also the next subsection). An additional finding was especially student teachers perceived PTL activities to enhance their self-knowledge as pre-service teachers and to support them in recognising their strengths, weaknesses, and areas for improvement.

Hypothesis 2: PTL activities become more effective when they encompass active learning activities, such as observing or analysing, rather than passive activities, such as listening

The findings showed that if (student) teachers participated actively, their professional development was promoted. More specifically, combined survey and observational data showed that they gained knowledge and implemented new ideas in classrooms (Koc, 2011; Zwart, Wubbels, Bolhuis, & Bergen, 2008). For example, in a collaborative teacher PTL activity, Meirink et al. (2007) showed in an extensive case study by triangulating inter-views and logbooks that, through active discussion, brainstorming, and reading, second-ary school teachers became aware of the strengths and weaknesses of their teaching methods, gained new ideas, and started to experiment with other teaching methods. This finding showed that being active (Hypothesis 2) related to reflection (Hypothesis 3). An example from a study on a coaching activity was that novice teachers reported their active participation in peer mentoring supported their confidence in their classroom management skills (Forbes, 2004). Schuck, who investigated online student teacher collab-oration, suggested that "the discussions encouraged reflection and understanding that might not otherwise have occurred" (2003, p. 24).

An additional finding was that active participation in collaborative activities was reported to reduce feelings of isolation (Hanewald & Gesthuizen, 2009; Harford & MacRuairc, 2008). This was revealed for both student teachers and in-service teachers.

Finally, in some studies evidence was found that combined the first and second hypoth-eses. These findings of survey as well as observational data showed that, during and/or after participating in coaching PTL activities, secondary school teachers began to exper-iment with implementing new ideas (Forbes, 2004; Zwart et al., 2008), and that student teachers applied newly gained knowledge in their field experience (Harford & MacRuairc, 2008; Lee & Baek, 2012). In other words, combining coaching activities and field experi-ence was successful.

Hypothesis 3: PTL activities become more effective when they have a reflection component

The findings showed that if reflection was elicited in PTL activities, for instance, by visiting another teacher's classroom (Hamilton, 2013; Manouchehri, 2001) or by working on colla-borative tasks (Lockhorst, Admiraal, & Pilot, 2010; Matthew, Felvegi, & Callaway, 2009), this led to a perceived gain in knowledge, self-knowledge, and ideas for improving one's own teaching skills. For example, after assessment PTL activities, student teachers reported reflection helped them identifying the strengths and weaknesses of their teaching and to consider ways to refine their teaching (Buchanan & Stern, 2012). Also, reflecting sup-ported student teachers to connect theory to practice (Jenkins, Garn, & Jenkins, 2005).

Hypothesis 4: PTL activities become more effective when they have a feedback component

Feedback is here considered as "information provided by an agent regarding aspects of one's performance or understanding" (Hattie & Timperley, 2007, p. 81). The findings showed that feedback was indeed a component of successful PTL activities, especially in teacher coaching activities. For instance, Zwart et al. (2008) demonstrated, by combining self-report data and observational data, that feedback among secondary school teachers influenced teachers' knowledge and self-knowledge. Novice teachers also reported that feedback supported them in developing effective classroom management techniques (Forbes, 2004). Similar findings were revealed for student teachers in coaching activities (Shin, Wilkins, & Ainsworth, 2007; So, 2012; Wilkins, Shin, & Ainsworth, 2009) and for student teachers in assessment activities (Buchanan, 2011). As an addition to the original hypothesis, student teachers in Wilkins et al.'s (2009) study reported that feedback processes enhanced their level of reflection.

Although only two studies with collaborative activities focused on feedback, their findings, both with less rigour, indicated feedback in collaborative settings enhances the outcomes of such activities. For instance, a dyad of secondary school teachers who were being followed by Manouchehri (2001) created their own observation and feedback protocol. After they had implemented this protocol, one of the teachers began to plan his instruction more deliberately.

As an addition to the original hypothesis, a third of the studies related to this hypothesis showed that feedback processes can be developed through interventions. For instance, Wilkins et al. (2009) showed the content of feedback changed throughout two semesters. In the first semester, feedback was directed at classroom management, delivery of lessons, and classroom climate. In the second semester, feedback became more directed at confidence in teaching. Apart from learning by doing, training peer assessment and feedback skills was effective for PTL. Sluijsmans et al. (2004), who compared groups of student teachers who had received training on peer assessment and groups that had not received training, demonstrated that trained student teachers gave more constructive feedback and applied more assessment criteria than the non-trained student teachers.

As another addition to the original hypothesis, two studies indicated challenges that arose while providing feedback in coaching PTL activities, for instance, a lack of constructive feedback and reflection being more technical (Shin et al., 2007). Other challenges were that student teachers perceived feedback was not always provided within a given time frame and that some student teachers sometimes preferred face-to-face feedback over online feedback because the former was considered more personal (Ruan & Beach, 2005).

Hypothesis 5: PTL activities become more effective when they apply video fragments of teaching

Findings suggested that applying video fragments can support the success of PTL activities, especially for student teachers. For instance, student teachers reported that video in coaching activities supported them in developing an understanding of teaching strategies and to reflect on which of these strategies were most appropriate (Bower, Cavanagh, Moloney, & Dao, 2011; So, 2012; So, Hung, & Yip, 2008). Other student teachers reported that video fragments supported connecting theory and practice (Baran & Cagiltay, 2010).

Furthermore, the ability of reviewing a video fragment was considered an advantage by participants and supported reflection (Amobi, 2005; Harford & MacRuairc, 2008).

At the same time, participants in about half of the studies related to this hypothesis reported disadvantages of using video fragments. These disadvantages were mostly more technical in nature, for instance, a limited camera view of the classroom (Forbes, 2004), lack of equipment, and lack of knowledge about video editing and exporting videos from editing software (Koc, 2011). Some student teachers as well as their pupils felt under scrutiny when they were videotaped (Harford & MacRuairc, 2008).

Only a few studies with in-service teachers applied video fragments in PTL. One of these studies showed that in the context of coaching teacher PTL activities, novice teachers liked that they had control over the camera's movement and its ability to zoom (Forbes, 2004).

Hypothesis 6: PTL activities become more effective when they are guided by a facilitator or group leader

The findings showed that, in contrast to the other two PTL activities, especially in colla-borative PTL activities, the presence of a facilitator or group leader was essential. For instance, the presence of such a leader was reported to positively affect collaborative activities (An, Kim, & Kim, 2008; Biasutti, 2011; Goos & Bennison, 2008), and lack thereof hindered collaborative activities (An et al., 2008; Fransen, Kirschner, & Erkens, 2011). To facilitate change, facilitators' most successful actions were to introduce ideas that are based on current levels of teacher understanding, interest, and skill (Borko et al., 1997) and to encourage sharing experiences (Goos & Bennison, 2008).

In coaching activities, such a facilitator was usually not present, probably because of its nature and the group size, which in coaching activities was smaller (two, three, or four par-ticipants) than in collaborative activities (mostly between four and nine participants). However, as an addition to the original hypothesis, coaching styles performed by the peer coaches seemed to matter. Engelen and Bergen (2002) demonstrated that a coopera-tive coaching style – wherein coach and coachee contributed evenly to discussing the observed lesson and brainstorming about possible improvements – during peer coaching was more appropriate for realising changed classroom behaviour of the coachee than a directive coaching style – wherein the coach contributed more to the discussion of the observed lesson and to possible improvement than the coachee. Moreover, students of secondary school teachers who had been coached in a cooperative coaching style reported more changed classroom behaviour than students of teachers who had been coached in a directive style (Engelen & Bergen, 2002). Moreover, Vincent and Jones (2008) speculated that, due to coaches, secondary school teachers were rapidly develop-ing their skills in using an interactive whiteboard. Student teachers in coaching activities said they did not need the instructor to take control over their coaching processes, especially when student teachers had to formulate feedback questions (Buchanan, 2011).

Hypothesis 7: PTL activities become more effective when they have a longer rather than shorter duration and follow-ups

Twenty-two studies had an intervention that lasted a course, which occurred mostly in student teacher interventions. Ten studies had an intervention that lasted one month or less; the shortest intervention was only 50 minutes. Five studies, mostly with student

teachers, lasted a semester. Ten studies, mostly with in-service teachers, lasted a (school) year, and a few more than one year. In seven studies, there was no duration reported.

Findings showed, regardless of the type of activity or the type of participants, a lack of time hindered the success of PTL activities (An et al., 2008; Forbes, 2004). Moreover, Borko et al. (1997) suggested time was needed for primary school teachers to change practice but could be a major obstacle because of their workload. Additionally, Amhag and Jacobsson (2009) showed by observational, longitudinal data that, over time, student teachers' collaboration reached deeper levels. Despite the fact that self-reported findings suggested a lack of time hindered the success of PTL activities, observational findings showed that when time was created, it was worth the effort.

We did not find evidence for the proposition that successful PTL activities arrange for follow-ups that are needed for transfer. There were simply no studies within our selection that included follow-up activities.

Contexts influencing PTL activities

Hypothesis 8: PTL activities become more effective when personal factors, such as motivation and beliefs, are taken into account

The findings showed that personal factors of participating (student) teachers influenced PTL activities. Motivation and prior knowledge and experience most especially were demonstrated to do so. Regarding motivation, the studies demonstrated that being motivated enhanced PTL activities (Buchanan & Stern, 2012; Donnelly, 2005; Hamilton, 2013; Helleve, 2007; Manouchehri, 2001; Meirink, 2007; Nicholas & Ng, 2009). For instance, student teachers who were committed to self-improvement reported to have welcomed constructive feedback from their peers (Buchanan & Stern, 2012).

Prior knowledge and experience (i.e., content and skill related) of individuals also affected the collaboration process. For instance, in-service teachers perceived communication skills (Main, 2012) and relationship-building skills (Hew & Hara, 2007) supported the collaboration process. Conversely, lack of experience with providing feedback (Shriki & Movshovitz-Hadar, 2011) or a lack of technical competence (Prestridge, 2009; see also the findings on Hypothesis 10) were perceived to hinder the collaboration process. At the same time, novice teachers in Forbes' (2004) study suggested that their lack of teaching experience did not influence their peer mentoring. Similar findings concerning prior knowledge and experience were also revealed for student teachers (Lamb, Lane, & Aldous, 2013; Ritchie & Peters, 2001; Schuck, 2003).

Furthermore, behaviour during the activities, expectations, perceptions, individual characteristics, and beliefs appeared to affect PTL activities. One example of behaviour that influenced collaboration PTL activities was that modelling desired mailing list behaviour, such as positive and supportive wording of messages (Hanewald & Gesthuizen, 2009), was perceived to positively influence other teacher participants' mailing list behaviours. Another example was found by Huang, Lubin, and Ge (2011), who showed that in a constructivist setting students demonstrated more active behaviours than students in a traditional setting, who were highly dependent on the instructor. The former students developed more creative products than the latter.

Expectations also appeared to influence the process. Manouchehri (2001) showed that when the collaboration process clashed with one of the teachers' expectations, this

teacher began to ask for more elaboration from her colleague, which he then provided. The novice teachers in Forbes' (2004) study had felt anxious about being observed and peer mentored; however, as a result of gaining experience with this activity, this anxiety decreased.

Perceptions during PTL activities were another personal factor that influenced PTL activities. For example, novice teachers felt anxious to coach each other in the beginning, but as they perceived benefits of the peer coaching, their anxiousness decreased (Forbes, 2004). Furthermore, student teachers felt they were not capable of assessing their peers, which might have hindered their peer assessment activities (Sluijsmans et al., 2004). Another study showed that if students felt online discussions did not meet their needs, they did not join in the discussion (Schuck, 2003).

Personality characteristics of individuals appeared to influence the PTL activities. For example, perfectionism (characteristic of individual) and the accompanying need to have firm control was perceived to hinder collaboration processes, especially if highly involved students had these characteristics (Eteläpelto et al., 2005).

Finally, a variety of teachers' beliefs seemed to influence the peer collaboration process. For instance, Pachler and Daly (2006) showed by self-reported data that teachers who believed in collaborative learning welcomed their own collaboration processes. This finding appeared in collaborative activities that were not initiated by policy (e.g., educational reforms) and shows that beliefs in themselves (e.g., not only educational beliefs) influenced the success of such activities.

Hypothesis 9: PTL activities become more effective when interpersonal factors, such as trust and respect, are taken into account

The findings showed that interpersonal factors influenced the success of PTL activities. Trust appeared essential for any activity (Biasutti, 2011; Estebaranz et al., 1999; Eteläpelto et al., 2005; Fransen et al., 2011; Helleve, 2007; Hew & Hara, 2007; Lamb et al., 2013; Shin et al., 2007; Wilkins et al., 2009). As Fransen et al. demonstrated: " … mutual trust appears to be conditional" (2011, p. 1109). Furthermore, sharing experiences was highly valued in a non-threatening space where all partners were perceived equal (Lamb et al., 2013; Ruan & Beach, 2005; So, 2012). Likewise, respecting each other, for instance, if a group member asked questions because they did not understand something, was demonstrated to support the collaboration (Roychoudhury & Roth, 1992). Not knowing each other was perceived to hinder collaboration processes (Nicholas & Ng, 2009). Challenges arose, for instance, if group members were perceived to be not committed to the tasks and their contribution was perceived unequal (Eteläpelto et al., 2005; Nicholas & Ng, 2009). This might have reduced trust.

Furthermore, communication between all group members in all possible directions was another, essential, interpersonal factor for PTL activities (Hanewald & Gesthuizen, 2009; Shriki & Movshovitz-Hadar, 2011). Additionally, combined findings from self-report and observational data showed that if group members had shared feelings of being one group working towards a shared goal, this enhanced success (Burron, James, & Ambrosio, 1993; Hew & Hara, 2007; Roychoudhury & Roth, 1992). Moreover, such a perceived collectivist orientation towards the team appeared positive, as without it participants felt isolated (Main, 2012).

For collaborative activities, collaboration skills affected the success of these processes. This means that group members determined who was responsible for what, and group members performed these tasks within a given time period (An et al., 2008; Baran & Cagiltay, 2010; Biasutti, 2011; Järvenoja & Järvelä, 2009). When this was the case, collaboration activities were enhanced. Group members participating and providing feedback to one another was also perceived to positively influence the process (An et al., 2008; Estebaranz et al., 1999). A prerequisite for collaboration was shared mental models (Fransen et al., 2011); however, to build and share such mental models, student teachers needed content knowledge (Hurme, Merenluoto, & Järvelä, 2009), and additionally, knowing each other supported the construction of shared mental models (Fransen et al., 2011). Perceiving a lack of collaborative skills in group members was shown to hinder the collaboration (An et al., 2008; Baran & Cagiltay, 2010; Järvenoja & Järvelä, 2009; Main, 2012; Shriki & Movshovitz-Hadar, 2011).

Online collaboration sometimes appeared challenging in the view of participants because online communication was hindered by its reliance on written text (An et al., 2008). Furthermore, communication was hindered by perceiving slow discussion processes during online collaboration (Nicholas & Ng, 2009).

Hypothesis 10: PTL activities become more effective when contextual factors, such as support from school leaders and ICT, are taken into account

The findings showed support from others in the environment (e.g., school or teacher education institute) was important for PTL activities, especially in the case of collaborative teacher activities and coaching activities. One example was support from the school organisation. Specifically, learning culture and administrative support were perceived to positively affect collaborative teacher activities (Main, 2012; Thijs & Van den Berg, 2002).

The findings also showed training certain skills was useful. For instance, training teachers on team learning prior to PTL was perceived to benefit collaboration activities (Main, 2012; Shriki & Movshovitz-Hadar, 2011). Furthermore, training student teachers in coaching activities was demonstrated to enhance the success of implementation of literacy teaching skills (Mallette, Maheady, & Harper, 1999).

The findings also demonstrated that if ICT was being used, it was important to make sure that the ICT was operable and that the participants, through training, experience, or support, learned how to operate the specific applications. Studies wherein online collaboration or coaching was applied showed that perceived technical problems (An et al., 2008; Forbes, 2004; Harford & MacRuairc, 2008; Schuck, 2003) and not being able to use the technology (Matthew et al., 2009; Shriki & Movshovitz-Hadar, 2011) hindered online PTL activities. Constant technical support in the form of a helpdesk was reported to solve most of the problems (Shriki & Movshovitz-Hadar, 2011). Additionally, Lockhorst et al. (2010), who implemented several online collaborative tasks over time, demonstrated that over the course of time, student teachers posted fewer messages that contained technological issues. In other words, to overcome technological issues, it was helpful to get to know how the technology works and how to work with it.

The findings revealed support for the influence of policy, especially for participating teachers. For example, in collaborative teacher learning activities, it was shown that if primary and secondary school teachers' beliefs cohered with the intended reform, they adapted their teaching practices according to the reform (Borko et al., 1997; Meirink,

2007). These findings showed educational beliefs interacted with activities initiated by policy. Similar results were found in studies with student teachers, yet they emerged from only a few studies. For instance, for collaborative student teacher activities, Baran and Cagiltay (2010) revealed that if reflection reports showed that students' goals corresponded with the intervention's goals, observations indicated that these students adapted more easily to the online communication system and participated more in the online activities. In other words, studies on student teacher peer learning activities also showed that educational beliefs that cohered with interventions led to more success.

Conclusions and discussion

The findings of the present realist synthesis provide an updated and extended overview of what makes peer (student) teacher learning (PTL) activities effective. We extended the scope on PTL by including studies on a *variety* of PTL activities. Because we used a typology of three activities (collaboration, coaching, and assessment) of about 20 different forms of PTL, we were able to explore whether there were differences in what makes these type of activities effective. Since the review was conducted via a grant and included only studies up to 2013, we have decided to link our major findings to some more recent studies in this section.

Because PTL activities are situated in a complex setting and influenced by a variety of processes and factors, we chose an approach to our literature review that accounted for this: realist synthesis (Pawson, 2002; Pawson et al., 2004). Realist synthesis supports answering the question of what works for whom under what circumstances. We explored what mechanisms (i.e., characteristics of the interventions) made PTL activities effective and which contexts (i.e., promoting or hindering factors) influenced this effectivity. This exploration consisted of developing and testing hypotheses. Table 2 summarises the findings of the testing of the hypotheses for each of the three different PTL activities – coaching, collaboration, and assessment – with two types of participants – student teachers and in-service teachers.

In conclusion, PTL activities need a focus on gaining subject matter and pedagogical content knowledge, with teachers actively working on gaining this knowledge. Such activities (or mechanisms) can be feedback and reflection, which enhance each other. Video recordings can also be used in the process of feedback and reflection. In the circumstance of collaborating PTL, a facilitator that guides the active involvement is necessary. In the circumstance of coaching PTL, a coach or coaching skills of the teachers themselves reinforce feedback and reflection. At the same time, motivation and prior knowledge that the participants bring to the PTL influence the process, and trust between participants is a prerequisite. Finally, circumstances within the organisation, such as support from the management and time, enable PTL, even though teacher motivation can compensate for the lack thereof.

Findings concerning the mechanisms showed that PTL activities should be focused on gaining subject matter and pedagogical content knowledge and should encompass activities that allow participants to actively work on this knowledge (cf. Van Veen et al., 2012). More recent studies confirm this finding (e.g., Carrillo, Maasen van den Brink, & Groot, 2016; Coenders & Verhoef, 2018; DeNeve & Devos, 2016; Kintz, Lane, Gotwals, & Cisterna, 2015). Furthermore, several studies suggested that feedback from peers elicits reflection

Table 2. Overview of findings of the hypotheses separately for the six settings.

Hypotheses	Coaching PTL		Collaborative PTL		Assessment PTL	
	Student teachers	In-service teachers	Student teachers	In-service teachers	Student teachers	In-service teachers
Mechanisms of the PTL activities						
Hypothesis 1	++	++	++	++	+	?
Self-knowledge: recognising strengths and weaknesses, area for improvement	+	?	+	?	+	?
Applying the newly gained knowledge	+	+	?	+	?	?
Hypothesis 2	+	++	++	++	?	?
Reduced feelings of isolation	?	?	+	+	?	?
Hypothesis 3	++	+	+	++	+	?
Hypothesis 4	++	++	+	+	+	?
Hypothesis 5	++	+	++	?	?	?
Technical issues with using video (camera)	+	+	+	?	?	?
Hypothesis 6	?	?	++	++	−	?
Style of coaching	+	++	?	?	?	?
Hypothesis 7	+	+	++	++	+	?
Follow-ups	?	?	?	?	?	?
Contexts influencing PTL activities						
Hypothesis 8						
Motivation/engagement	+	++	++	++	+	?
Prior knowledge and skills	+	+	++	++	?	?
Behaviour	+	?	++	+	?	?
Expectations	?	+	+	+	?	?
Perceptions	+	+	+	+	+	?
Individual characteristics	?	?	+	?	?	?
Beliefs	?	++	+	++	?	?
Hypothesis 9						
Trust	++	+	++	++	?	?
Communication	+	+	++	++	?	?
Collaboration skills	?	?	++	++	?	?
Hypothesis 10						
Support from environment	+	+	?	+	?	?
ICT	++	+	++	+	?	?
Training	+	?	?	+	+	?
Policy	?	?	+	+	?	?

Note: + = there were clues of evidence; ++ = evidence was clear; ? = there were no findings supporting the hypotheses; − = clues of evidence rejected hypothesis.

on the part of the feedback receiver (Showers & Joyce, 1996) and, moreover, that feedback and reflection were important characteristics of effective PTL activities. Again, more recent studies confirm this finding (e.g., Coenders & Verhoef, 2018; DeNeve & Devos, 2016). Additionally, a facilitator is needed, especially in collaborative activities. Even though the included studies remained unclear about what a facilitator should do and know in order to enhance collaboration, recent studies have begun to explore this role into more detail. For example, Borko, Koellner, and Jacobs (2014) examined novice facilitators' behaviour in their Problem Solving Cycle. They found that these facilitators were succeeding in developing respect and relationships in their groups and in choosing video fragments to discuss within the group and, at the same time, that these facilitators were less successful in guiding discussions in such a way that developing subject matter and pedagogical content knowledge was enhanced. Another example is a study by Binkhorst, Poortman, McKenney, and Van Joolingen (2018), who explored how vertical and horizontal leadership influenced collaboration processes in teacher design teams, and showed that

different phases in the collaboration needed either vertical leadership (e.g., in planning) or horizontal leadership (e.g., in brainstorming). A third example is a study by An, Shin, and Lim (2009), who implemented three different facilitation strategies in online student teacher collaboration. Their findings showed that in each group the interaction evolved in a different way. For instance, when the facilitator responded to each student teacher's messages and invited others to reply, the interaction between student teachers was not so intense as in the group where the facilitator did not respond to individual student teachers' messages but did strongly encourage them to reply to each other's' messages. Future reviews can include the role of the facilitator into more detail.

Findings showed that while video excerpts can be used in PTL activities to enhance, for example, bridging theory and practice (cf. Brophy, 2003), participants also need to learn to operate a camera. The recent study of Coenders and Verhoef (2018) showed that teachers involved in Lesson Study learned from observing each other's lessons. Moreover, time is an important mechanism, as activities need to span a certain period in order to make learning possible (cf. Desimone, 2009; Lauer et al., 2014; Van Veen et al., 2012). Even though it remained unclear from the included studies how much time is needed, a recent review showed that the longer the length of training in professional development programmes is, the more likely teachers' development of their pedagogical content knowledge is enhanced and, moreover, student outcomes in mathematics (Carrillo et al., 2016).

Summarising the mechanisms, we conclude that PTL activities should span at least one school year, during which participants are actively involved, using reflection and feedback processes which can be enhanced by observing lessons, in order to gain subject matter and pedagogical content knowledge. Facilitators are needed to promote participants' active involvement and reflection and feedback processes.

Findings concerning contexts influencing the success of PTL activities showed a large variety of factors. The personal factors motivation (cf. Ryan & Deci, 2000) and having some prior knowledge and/or skills to build on (cf. Hew & Hara, 2007; Main, 2012) were addressed as important for the effectiveness of PTL activities in the majority of studies. A more recent study showed the importance of self-efficacy in novice teachers' professional learning communities, aiming to implement differentiated instruction (DeNeve, Devos, & Tuytens, 2015). The interpersonal factors trust (cf. Salas, Sims, & Burke, 2005; Van Gennip et al., 2009) and collaboration skills (cf. Salas et al., 2005; Van den Bossche, Gijselaers, Segers, & Kirschner, 2006) were considered important for the effectiveness of PTL activities in the majority of studies. More recent studies confirm these insights (e.g., Raes, Kyndt, Decuyper, Van den Bossche, & Dochy, 2015; Wang, 2015). School factors, namely, support from the management, time, and training (cf. Avalos, 2011; Van Veen et al., 2012) were addressed as important for the effectiveness of PTL activities in the majority of studies. Again, recent studies confirmed these insights (e.g., DeNeve & Devos, 2016, 2017; DeNeve et al., 2015; Hadfield & Jopling, 2016; Wang, 2015). At the same time, studies departing from an organisational perspective rather than a learning activities perspective add to and nuance this view. For example, Penuel, Riel, Krause, and Frank (2009) used the social capital theory, and Daly, Moolenaar, Bolivar, and Burke (2010) used social networking theory as lenses to examine teacher-to-teacher interactions in the context of curricular changes. Both studies showed how school leaders act as gatekeeper between the imposed changes and the practice in schools. Daly et al. (2010) added that high interdependency in teacher networks strengthened ownership and empowerment of teachers.

The findings concerning mechanisms and factors influencing the success of PTL activities resonate with Fishbein and Ajzen's theory of planned behaviour (2010), which postulates that proximal and distal factors influence behaviour. Proximal factors are closer to behaviour, in our case (for example) motivation and collaboration skills. Distal factors are at a longer distance, in our case (for example) the characteristics of PTL activities and school processes. Fishbein and Ajzen suggested that proximal factors mediate between distal factors and behaviour. None of the included studies explicitly connected proximal and distal factors. Therefore, even though we tried to delve into the black box of these processes, the body of knowledge remains superficial, and questions for future research therefore could be: Given a well-designed PTL activity, how do teachers use this activity and how does their motivation play a role? Given a well-designed PTL activity, how do teachers collaborate or coach each other? Which personal or interpersonal factors can compensate for a PTL activity lacking a given mechanism? The recent studies of DeNeve and colleagues (DeNeve & Devos, 2016, 2017; DeNeve et al., 2015), wherein they focus on the interplay of school leadership, perceived characteristics of teacher professional learning communities, and teacher self-efficacy, exemplify how these kind of questions can be explored in both quantitative and qualitative approaches.

While the findings confirm former insights, at the same time the quality and types of study were questionable. Borko (2004) distinguished different types of research (a) Phase 1 studies focusing on one professional development programme in one site, (b) Phase 2 studies focusing on one professional development programme in different sites, and (c) Phase 3 studies focusing one different professional development programmes in different sites (Borko, 2004, p. 4). The majority of the included studies can be characterised as Phase 1 studies, as they tended to focus on only a few mechanisms and contexts. In alignment with Borko (2004), we recommend future studies to move beyond single case studies and beyond mapping only a few mechanisms and contexts so that they account for the complex setting that PTL is. Furthermore, while interpreting the findings, we experienced that the first and second hypotheses were related; the same held true for the third and fourth hypotheses. Therefore, it remained unclear which specific characteristics of PTL make it effective in terms of outcomes. Another limitation of our realist synthesis was that in some studies the rigour of methods needed to draw conclusions (Pawson et al., 2004, 2005) was questionable. This came to the fore in terms of the many self-reported instruments that were used in most of the studies (see also Appendix 1; see, e.g., Lawless & Pellegrino, 2007; Thurlings & Den Brok, 2017; Vescio et al., 2008). Therefore, for future research it is recommended to rely less on self-reported data. Additionally, in several publications, the methodology sections lacked certain information, such as clear descriptions of scoring of interviews and observations (cf. former literature studies such as Lawless and Pellegrino, 2007; and Vescio et al., 2008). Furthermore, in about half of the articles, it was not mentioned at which type of school (primary or secondary) participants were based, nor was it explained what the goals of the interventions studied were. Another limitation might be the operationalisation of what constitutes "peers". Many studies were excluded based on the peer criterion, simply because the publications did not give enough information about the equality of status within the PTL. As we only included three relevant publications on peer assessment activities, which were all restricted to student teachers, it remains

Table 3. Context-mechanisms-outcomes configurations.

For developers of interventions	For school management and teacher educators	For teachers and student teachers
Content and activities		
Implement active participation through, e.g., discussions, visiting classrooms, and brainstorming	Implement active participation through, e.g., discussions, visiting classrooms, and brainstorming	Be prepared to actively participate and to try out new teaching methods within your classroom that suit your interests
Focus on content knowledge, pedagogical content knowledge, or teaching skills that have or meet teachers' interests	Focus on content knowledge, pedagogical content knowledge or teaching skills that have or meet teachers' interests	
Combine activities wherein participants develop new knowledge or teaching skills with the application in their classrooms (for teachers) and their field experience (for student teachers)	Combine activities wherein participants develop new knowledge or teaching skills with the application in their classrooms (for teachers) and their field experience (for student teachers)	
Support participants in developing ideas for implementation of knowledge	Support participants in developing ideas for implementation of knowledge	
Facilitate the application of new knowledge through experimentation	Facilitate the application of new knowledge through experimentation	
	Support student teachers to develop self-knowledge about their strengths, weaknesses, and areas for improvement	
Pre-condition: time		
	Make sure participants know that they need to invest time for participating in PTL activities and to achieve outcomes	Be prepared and willing to invest time
	Create time, so that participants can simultaneously work together	
	Create interventions that last at least one school year	
Pre-condition: video/ICT		
When using technological applications, make sure that these are operable and that participants can use these applications	When using technological applications, make sure that these are operable and that participants can use these applications	If you use any technological application, be prepared to learn to deal with it.
Provide training and/or constant support on participants' skills in using technological applications	Provide training and/or constant support on participants' skills in using technological applications	
Realise that video is a powerful *tool*	Realise that video is a powerful *tool*	Realise that video is a powerful *tool*
	When using video recordings, train participants in installing and using equipment	When using video recordings, train yourself to use the equipment (if applicable)
		Using video recordings may seem threatening but is a powerful tool
		Respect your colleagues who open up their classrooms by showing videos; if you are showing videos, try to prevent being hesitant: Using it more often helps
Pre-condition: training		
Provide participants with training on necessary skills	Provide participants with training on necessary skills	Ask for training if you feel you need training to successfully participate
Pre-condition: support		
	School management and teacher educators should provide support, in	Ask your school leader or teacher educator for the needed support

Table 3. Continued.

For developers of interventions	For school management and teacher educators	For teachers and student teachers
	terms of necessary resources as well as social support	
Personal		
Align interventions with participants' beliefs	If participants' beliefs do not cohere with the intervention, do not expect to yield intended outcomes Be aware of and articulate participants' beliefs	Select to participate in activities that cohere with your beliefs about teaching Behave pro-actively, communicate your expectations and perceptions especially if these clash with the intervention
Align interventions with participants' prior knowledge and experiences	Be aware of participants' behaviour, expectations, perceptions, and characters. As a facilitator, coach, or group member signal and discuss hindering behaviour, expectations, perceptions, and characters	Choose to participate in interventions that build on your knowledge and experiences
Working together		
	When using online collaboration or coaching, be aware of its reliance on written text and the lack of non-verbal behaviour. Knowing the group members beforehand and/or organising face-to-face meetings might be necessary	When using online collaboration or coaching, be aware of its reliance on written text and the lack of non-verbal behaviour
	Build trust between participants	Get to know your co-participants Build trust and be trustworthy
	Communicate and monitor work processes Work on collaborative skills, especially in collaborative activities Formulate a shared goal, especially in collaborative activities	Communicate and monitor work processes Work on collaborative skills, especially in collaborative activities Formulate a shared goal, especially in collaborative activities
Facilitator		
	Add a facilitator to collaborative groups, who introduces new ideas to the group and encourages active participation Work on coaching skills, through employing skilled coaches or training participants	Work on your coaching skills
Feedback during activities		
Develop tasks or activities that elicit providing feedback		When you are providing feedback, try to be constructive in your remarks
Organise feedback moments, for example, using videos	Organise feedback moments, for example, using videos	When you are receiving feedback, ask the provider(s) to be as constructive as possible
Support participants' competence in providing constructive feedback	Support participants' competence in providing constructive feedback	
Reflection during and after activities		
Organise reflection moments, wherein participants discuss their reflections	Organise reflection moments, wherein participants discuss their reflections	Reflect on your experiences within the PTL activity and be prepared to share these with your co-participants
Elicit reflection, for instance, by visiting each other's classrooms or by organising feedback moments	Elicit reflection, for instance, by visiting each other's classrooms or by organising feedback moments	

unclear if and how findings are the same for such assessment activities. Future studies on peer assessment are therefore needed. Afterwards, when sufficient studies have been conducted, our recent realist synthesis could be updated and expanded with peer assessment PTL activities. Another issue can be that some mechanisms remain hidden. In the data extraction, we solely used the published articles. For example, a few studies indicated that personal factors such as beliefs and attitudes influence the process of PTL; however, it remained unclear how, which is an enormous challenge for future studies. A final limitation is that the recent realist synthesis mainly focused on mechanisms and contexts or factors, and did (not as other realist syntheses tend to do) use outcomes, sometimes at different levels, as a starting point. The recent realist synthesis focused on "effective PTL activities"; there were no studies that demonstrated "ineffective PTL activities" (cf. Jones, Shipman, & Ogrinc, 2014). Future literature reviews might explore how, for example, teachers' subject matter knowledge or their classroom management skills or student achievement are enhanced through specific mechanisms or types of PTL activities.

In conclusion, findings confirmed insights concerning effective student teacher learning and teacher professional development and revealed what characteristics make PTL effective (Ackland, 1991; Avalos, 2011; Capps et al., 2012; Desimone, 2009; Evers et al., 2011; Lauer et al., 2014; Salas et al., 2005; Van Gennip et al., 2009; Vangrieken et al., 2015; Van Veen et al., 2012; Vescio et al., 2008), especially for coaching and for collaboration activities.

On the basis of the findings, we formulated a number of practical implications that can be used to evaluated PTL activities, summarised into a checklist (Table 3). These can be seen as context-mechanism-outcome configurations (Dalkin, 2014; Wong, Westhorp, et al., 2013). These practical implications and the checklist were presented and shared during a symposium with stakeholders (see section Stakeholders). The need for time and resources (Hypotheses 7 and 10) as well as the need for support from the school organisation (Hypothesis 10) were recognised as a necessity for effective PTL activities. Additionally, attendees of the symposium confirmed that PTL activities should span over at least one year to facilitate actual change. Furthermore, being critical friends (Hypotheses 2, 3, and 4) and sharing a goal and process (Hypothesis 9) were considered important. Moreover, the attendees stressed that contexts are influential for making PTL effective, such as the level of education (i.e., primary or secondary education). In other words, they suggested that the circumstances wherein PTL takes place differ (e.g., primary vs. secondary education, country, policy-based intervention vs. non-policy-based intervention, the amount of experience or expertise of participants) and that each circumstance has its own unique characteristics that could be influential. Consequently, research findings need to be translated to the specific circumstances; there is no one-size-fits-all panacea (see also Raes et al., 2015; Wong, Westhorp, et al., 2013). At the same time, researchers exploring specific circumstances need to elaborate more about these circumstances in their publications to support practice and research.

A final contribution of this literature review is the application of realist synthesis in educational research. Although realist synthesis was originally developed for health care interventions, Pawson et al. (2004) suggested that the approach would also be applicable to other fields and disciplines, such as education. We have experienced and shown that

realist synthesis can indeed be applied to education, and that it yields practical, relevant implications for practice. Realist synthesis aims not only to review what works for whom under what circumstances, but also supports the quest for evidence-based practice (Pawson, 2002; Pawson et al., 2004). Evidence-based practice is an emerging aspect in education (e.g., Hew & Cheung, 2013), and this even more so calls for conducting realist syntheses on other interventions in education.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Notes on contributors

Marieke Thurlings works as an assistant professor at Eindhoven School of Education. Her expertise lies in teacher professional development, STEM education, and systematic literature reviews. She has published in several journals, such as *Review of Educational Research* and *Australasian Journal of Educational Technology*.

Perry den Brok is full professor in education and learning sciences at Wageningen University & Research. His research interests are related to topics such as learning environments, teacher learning and professional development, higher education, and teacher–student interpersonal relationships. He has edited several books and written many articles on these topics.

References

References marked with an asterisk indicate studies included in the review.

Ackland, R. (1991). A review of the peer coaching literature. *Journal of Staff Development*, *12*(1), 22–27.

*Amhag, L., & Jakobsson, A. (2009). Collaborative learning as a collective competence when students use the potential of meaning in asynchronous dialogues. *Computers & Education*, *52*(3), 656–667. doi:10.1016/j.compedu.2008.11.012

*Amobi, F. A. (2005). Preservice teachers' reflectivity on the sequence and consequences of teaching actions in a microteaching experience. *Teacher Education Quarterly*, *32*(1), 115–130.

*An, H., Kim, S., & Kim, B. (2008). Teacher perspectives on online collaborative learning: Factors perceived as facilitating and impeding successful online group work. *Contemporary Issues in Technology and Teacher Education*, *8*(1), 65–83. Retrieved from http://www.citejournal.org/volume-8/issue-1-08/general/teacher-perspectives-on-online-collaborative-learning-factors-perceived-as-facilitating-and-impeding-successful-online-group-work

An, H., & Shin, S., & Lim, K. (2009). The effects of different instructor facilitation approaches on students' interactions during asynchronous online discussions. *Computers & Education*, *53*(3), 749–760. doi:10.1016/j.compedu.2009.04.015

*Arvaja, M. (2007). Contextual perspective in analysing collaborative knowledge construction of two small groups in web-based discussion. *International Journal of Computer-Supported Collaborative Learning*, *2*(2–3), 133–158. doi:10.1007/s11412-007-9013-5

Avalos, B. (2011). Teacher professional development in *Teaching and Teacher Education* over ten years. *Teaching and Teacher Education*, *27*(1), 10–20. doi:10.1016/j.tate.2010.08.007

*Baran, B., & Cagiltay, K. (2010). The dynamics of online communities in the activity theory framework. *Educational Technology & Society*, *13*(4), 155–166.

*Biasutti, M. (2011). The student experience of a collaborative e-learning university module. *Computers & Education*, *57*(3), 1865–1875. doi:10.1016/j.compedu.2011.04.006

Binkhorst, F., Poortman, C. L., McKenney, S. E., & Van Joolingen, W. R. (2018). Revealing the balancing act of vertical and shared leadership in Teacher Design Teams. *Teaching and Teacher Education*, *72*, 1–12. doi:10.1016/j.tate.2018.02.006

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, *33*(8), 3–15. doi:10.3102/0013189X033008003

Borko, H., Koellner, K., & Jacobs, J. (2011). Meeting the challenges of scale: The importance of preparing professional development leaders. *Teachers College Record*. http://www.tcrecord.org (ID Number: 16358)

Borko, H., Koellner, K., & Jacobs, J. (2014). Examining novice teacher leaders' facilitation of mathematics professional development. *The Journal of Mathematical Behavior*, *33*, 149–167. doi:10.1016/j.jmathb.2013.11.003

*Borko, H., Mayfield, V., Marion, S., Flexer, R., & Cumbo, K. (1997). Teachers' developing ideas and practices about mathematics performance assessment: Successes, stumbling blocks, and implications for professional development. *Teaching and Teacher Education*, *13*(3), 259–278. doi:10.1016/S0742-051X(96)00024-8

*Bower, M., Cavanagh, M., Moloney, R., & Dao, M. (2011). Developing communication competence using an online Video Reflection system: Pre-service teachers' experiences. *Asia-Pacific Journal of Teacher Education*, *39*(4), 311–326. doi:10.1080/1359866X.2011.614685

*Britton, L. R., & Anderson, K. A. (2010). Peer coaching and pre-service teachers: Examining an underutilised concept. *Teaching and Teacher Education*, *26*(2), 306–314. doi:10.1016/j.tate.2009.03.008

Brophy, J. (Ed.). (2003). *Using video in teacher education (Advances in research on teaching, Vol. 10)*. Amsterdam: Elsevier.

*Buchanan, M. T. (2011). Peer reviewing preservice teachers of Christian higher education. *Christian Higher Education*, *10*(5), 464–481. doi:10.1080/15363759.2010.515455

*Buchanan, M. T., & Stern, J. (2012). Pre-service teachers' perceptions of the benefits of peer review. *Journal of Education for Teaching*, *38*(1), 37–49. doi:10.1080/02607476.2012.643654

*Burron, B., James, M. L., & Ambrosio, A. L. (1993). The effects of cooperative learning in a physical science course for elementary/middle level preservice teachers. *Journal of Research in Science Teaching*, *30*(7), 697–707. doi:10.1002/tea.3660300708

Capps, D. K., Crawford, B. A., & Constas, M. A. (2012). A review of empirical literature on inquiry professional development: Alignment with best practices and a critique of the findings. *Journal of Science Teacher Education*, *23*(3), 291–318. doi:10.1007/s10972-012-9275-2

Carrillo, C., Maasen van den Brink, H., & Groot, W. (2016). *Professional development programs and their effects on student achievement: A systematic review of evidence* (TIER Working Paper No. 16/03). Maastricht: TIER, Maastricht University.

*Christ, T., Arya, P., & Chiu, M. M. (2012). Collaborative peer video analysis: Insights about literacy assessment and instruction. *Journal of Literacy Research*, *44*(2), 171–199. doi:10.1177/1086296X12440429

*Clarke, M. (2009). The discursive construction of interpersonal relations in an online community of practice. *Journal of Pragmatics*, *41*(11), 2333–2344. doi:10.1016/j.pragma.2009.04.001

Coenders, F., & Verhoef, N. (2018). Lesson Study: Professional development (PD) for beginning and experienced teachers. *Professional Development in Education*. Advance online publication. doi:10.1080/19415257.2018.1430050

Dalkin, S. (2014). *The realist evaluation of a palliative integrated care pathway in primary care: What works, for whom and in what circumstances?* (Doctoral dissertation). Retrieved from http://nrl.northumbria.ac.uk/18251/

Dalkin, S., Greenhalgh, J., Jones, D., Cunningham, B., & Lhussier, M. (2015). What's in a mechanism? Development of a key concept in realist evaluation. *Implementation Science*, *10*:49. doi:10.1186/s13012-015-0237-x

Daly, A. J., Moolenaar, N. M., Bolivar, J. M., & Burke, P. (2010). Relationships in reform: The role of tea-chers' social networks. *Journal of Educational Administration*, *48*(3), 359–391. doi:10.1108/09578231011041062

Darling-Hammond, L., & Snyder, J. (2000). Authentic assessment of teaching in context. *Teaching and Teacher Education*, *16*(5–6), 523–545. doi:10.1016/S0742-051X(00)00015-9

DeNeve, D., & Devos, G. (2016). The role of environmental factors in beginning teachers' professional learning related to differentiated instruction. *School Effectiveness and School Improvement*, *27*(4), 357–379. doi:10.1080/09243453.2015.1122637

DeNeve, D., & Devos, G. (2017). How do professional learning communities aid and hamper pro-fessional learning of beginning teachers related to differentiated instruction? *Teachers and Teaching*, *23*(3), 262–283. doi:10.1080/13540602.2016.1206524

DeNeve, D., Devos, G., & Tuytens, M. (2015). The importance of job resources and self-efficacy for beginning teachers' professional learning in differentiated instruction. *Teaching and Teacher Education*, *47*, 30–41. doi:10.1016/j.tate.2014.12.003

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, *38*(3), 181–199. doi:10.3102/0013189X08331140

Dochy, F., Segers, M. S. R., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, *24*(3), 331–350. doi:10.1080/03075079912331379935

*Donnelly, A. M. (2005). Let me show you my portfolio! Demonstrating competence through peer interviews. *Action in Teacher Education*, *27*(3), 55–63. doi:10.1080/01626620.2005.10463390

*El-Deghaidy, H., & Nouby, A. (2008). Effectiveness of a blended e-learning cooperative approach in an Egyptian teacher education programme. *Computers & Education*, *51*(3), 988–1006. doi:10.1016/j.compedu.2007.10.001

*Engelen, A., & Bergen, T. C. M. (2002). Op zoek naar factoren die van coaching een krachtige leer-omgeving voor docenten maken [Indicators of successful peer coaching in powerful learning environments]. *Pedagogische Studiën*, *79*(4), 305–321.

*Estebaranz, A., Mingorance, P., & Marcelo, C. (1999). Teachers' work groups as professional develop-ment: What do the teachers learn? *Teachers and Teaching*, *5*(2), 153–169. doi:10.1080/1354060990050202

*Eteläpelto, A., Littleton, K., Lahti, J., & Wirtanen, S. (2005). Students' accounts of their participation in an intensive long-term learning community. *International Journal of Educational Research*, *43*(3), 183–207. doi:10.1016/j.ijer.2006.06.011

Evers, A. T., Kreijns, K., Van der Heijden, B. I. J. M., & Gerrichhauzen, J. T. G. (2011). An organizational and task perspective model aimed at enhancing teachers' professional development and occu-pational expertise. *Human Resource Development Review*, *10*(2), 151–179. doi:10.1177/1534484310397852

Fishbein, M. & Ajzen, I. (2010). *Predicting and changing behavior: The reasoned action approach*. New York, NY: Psychology Press.

*Forbes, C. T. (2004). Peer mentoring in the development of beginning secondary science teachers: Three case studies. *Mentoring and Tutoring*, *12*(2), 219–239. doi:10.1080/1361126042000239956

*Fransen, J., Kirschner, P. A., & Erkens, G. (2011). Mediating team effectiveness in the context of col-laborative learning: The importance of team and task awareness. *Computers in Human Behavior*, *27*(3), 1103–1113. doi:10.1016/j.chb.2010.05.017

*Fry, J. M., & Hin, M. K. T. (2006). Peer coaching with interactive wireless technology between student teachers: Satisfaction with role and communication. *Interactive Learning Environments*, *14*(3), 193–204. doi:10.1080/10494820600852969

*Goos, M. E., & Bennison, A. (2008). Developing a communal identity as beginning teachers of math-ematics: Emergence of an online community of practice. *Journal of Mathematics Teacher Education*, *11*(1), 41–60. doi:10.1007/s10857-007-9061-9

*Grion, V., & Varisco, B. M. (2007). On line collaboration for building a teacher professional identity. *PsychNology Journal*, *5*(3), 271–284.

Hadfield, M., & Jopling, M. (2016). Problematizing lesson study and its impacts: Studying a highly contextualized approach to professional learning. *Teaching and Teacher Education*, *60*, 203–214. doi:10.1016/j.tate.2016.08.001

*Hamilton, E. R. (2013). His ideas are in my head: Peer-to-peer teacher observations as professional development. *Professional Development in Education*, *39*(1), 42–64. doi:10.1080/19415257.2012. 726202

*Hanewald, R., & Gesthuizen, R. (2009). Sustainability in an online community of practice: The case study of a group of secondary school educators in Victoria. *The Australian Journal of Teacher Education*, *34*(5), 26–42. doi:10.14221/ajte.2009v34n5.3

*Harford, J., & MacRuairc, G. (2008). Engaging student teachers in meaningful reflective practice. *Teaching and Teacher Education*, *24*(7), 1884–1892. doi:10.1016/j.tate.2008.02.010

*Harrington, H. L., & Hathaway, R. S. (1994). Computer conferencing, critical reflection, and teacher development. *Teaching and Teacher Education*, *10*(5), 543–554. doi:10.1016/0742-051X (94)90005-1

*Harrington, H. L., & Quinn-Leering, K. (1996). Considering teaching's consequences. *Teaching and Teacher Education*, *12*(6), 591–607. doi:10.1016/S0742-051X(96)00007-8

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. doi:10.3102/003465430298487

*Helleve, I. (2007). In an ICT-based teacher-education context: Why was our group "the magic group"? *European Journal of Teacher Education*, *30*(3), 267–284. doi:10.1080/02619760701486118

Hew, K. F., & Cheung, W. S. (2013). Use of Web 2.0 technologies in K-12 and higher education: The search for evidence-based practice. *Educational Research Review*, *9*, 47–64. doi:10.1016/j.edurev. 2012.08.001

*Hew, K. F., & Hara, N. (2007). Empirical study of motivators and barriers of teacher online knowledge sharing. *Educational Technology Research and Development*, *55*(6), 573–595. doi:10.1007/s11423-007-9049-2

*Howard, C. D., Barrett, A. F., & Frick, T. W. (2010). Anonymity to promote peer feedback: Pre-service teachers' comments in asynchronous computer-mediated communication. *Journal of Educational Computing Research*, *43*(1), 89–112. doi:10.2190/EC.43.1.f

*Huang, K., Lubin, I. A., & Ge, X. (2011). Situated learning in an educational technology course for pre-service teachers. *Teaching and Teacher Education*, *27*(8), 1200–1212. doi:10.1016/j.tate.2011. 06.006

*Hurme, T.-R., Merenluoto, K., & Järvelä, S. (2009). Socially shared metacognition of pre-service primary teachers in a computer-supported mathematics course and their feelings of task difficulty: A case study. *Educational Research and Evaluation*, *15*(5), 503–524. doi:10.1080/13803610903444659

*Järvenoja, H., & Järvelä, S. (2009). Emotion control in collaborative learning situations: Do students regulate emotions evoked by social challenges? *The British Journal of Educational Psychology*, *79*(3), 463–481. doi:10.1348/000709909X402811

*Jenkins, J. M., Garn, A., & Jenkins, P. (2005). Preservice teacher observations in peer coaching. *Journal of Teaching in Physical Education*, *24*(1), 2–23. doi:10.1123/jtpe.24.1.2

*Jenkins, J. M., & Veal, M. L. (2002). Preservice teachers' PCK development during peer coaching. *Journal of Teaching in Physical Education*, *22*(1), 49–68. doi:10.1123/jtpe.22.1.49

Jones, A. C., Shipman, S. A., & Ogrinc, G. (2014). Key characteristics of successful quality improvement curricula in physician education: A realist review. *BMJ Quality and Safety*, *24*(1), 77–88. doi:10.1136/bmjqs-2014-002846

Kintz, T., Lane, J., Gotwals, A., & Cisterna, D. (2015). Professional development at the local level: Necessary and sufficient conditions for critical colleagueship. *Teaching and Teacher Education*, *51*, 121–136.doi:10.1016/j.tate.2015.06.004

*Koc, M. (2011). Let's make a movie: Investigating pre-service teachers' reflections on using video-recorded role playing cases in Turkey. *Teaching and Teacher Education*, *27*(1), 95–106. doi:10. 1016/j.tate.2010.07.006

*Lamb, P., Lane, K., & Aldous, D. (2013). Enhancing the spaces of reflection: A buddy peer-review process within physical education initial teacher education. *European Physical Education Review*, *19*(1), 21–38. doi:10.1177/1356336X12457293

Lauer, P. A., Christopher, D. E., Firpo-Triplett, R., & Buchting, F. (2014). The impact of short-term professional development on participant outcomes: A review of the literature. *Professional Development in Education*, *40*(2), 207–227. doi:10.1080/19415257.2013.776619

Lawless, K. A., & Pellegrino, J. W. (2007). Professional development in integrating technology into teaching and learning: Knowns, unknowns, and ways to pursue better questions and answers. *Review of Educational Research*, *77*(4), 575–614. doi:10.3102/0034654307309921

*Lee, H.-J., & Baek, E. (2012). Facilitating deep learning in a learning community. *International Journal of Technology and Human Interaction*, *8*(1), 1–13. doi:10.4018/jthi.2012010101

Little, J. W. (1990). The persistence of privacy: Autonomy and initiative in teachers' professional relations. *Teachers College Record*, *91*(4), 509–536.

*Lockhorst, D., Admiraal, W., & Pilot, A. (2010). CSCL in teacher training: What learning tasks lead to collaboration? *Technology, Pedagogy and Education*, *19*(1), 63–78. doi:10.1080/1475939090 3579190

Lu, H.-L. (2010). Research on peer coaching in preservice teacher education: A review of literature. *Teaching and Teacher Education*, *26*(4), 748–753. doi:10.1016/j.tate.2009.10.015

*Main, K. (2012). Effective middle school teacher teams: A ternary model of interdependency rather than a catch phrase. *Teachers and Teaching*, *18*(1), 75–88. doi:10.1080/13540602.2011.622556

*Mallette, B., Maheady, L., & Harper, G. F. (1999). The effects of reciprocal peer coaching on preservice general educators' instruction of students with special learning needs. *Teacher Education and Special Education*, *22*(4), 201–216. doi:10.1177/088840649902200402

*Manouchehri, A. (2001). Collegial interaction and reflective practice. *Action in Teacher Education*, *22*(4), 86–97. doi:10.1080/01626620.2001.10463032

*Matthew, K. I., Felvegi, E., & Callaway, R. A. (2009). Wiki as a collaborative learning tool in a language arts methods class. *Journal of Research on Technology in Education*, *42*(1), 51–72. doi:10.1080/ 15391523.2009.10782541

*Meirink, J. A. (2007). *Individual teacher learning in a context of collaboration in teams* (Unpublished doctoral dissertation). Universiteit Leiden, Leiden, The Netherlands.

*Meirink, J. A., Imants, J., Meijer, P. C., & Verloop, N. (2010). Teacher learning and collaboration in innovative teams. *Cambridge Journal of Education*, *40*(2), 161–181. doi:10.1080/0305764X.2010.481256

*Meirink, J. A., Meijer, P. C., & Verloop, N. (2007). A closer look at teachers' individual learning in collaborative settings. *Teachers and Teaching*, *13*(2), 145–164. doi:10.1080/13540600601152496

*Meirink, J. A., Meijer, P. C., Verloop, N., & Bergen, T. C. M. (2009). How do teachers learn in the workplace? An examination of teacher learning activities. *European Journal of Teacher Education*, *32*(3), 209–224. doi:10.1080/02619760802624096

*Nicholas, H., & Ng, W. (2009). Fostering online social construction of science knowledge with primary pre-service teachers working in virtual teams. *Asia-Pacific Journal of Teacher Education*, *37*(4), 379–398. doi:10.1080/13598660903050336

*Oh, S., & Jonassen, D. H. (2007). Scaffolding online argumentation during problem solving. *Journal of Computer Assisted Learning*, *23*(2), 95–110. doi:10.1111/j.1365-2729.2006.00206.x

*Olson, J. K. (2007). Preservice teachers' thinking within a research-based framework: What informs decisions? *International Journal of Science and Mathematics Education*, *5*(1), 49–83. doi:10.1007/ s10763-005-9027-2

*Pachler, N., & Daly, C. (2006). Professional teacher learning in virtual environments. *E-Learning*, *3*(1), 62–74. doi:10.2304/elea.2006.3.1.62

Pawson, R. (2002). Evidence-based policy: The promise of "realist synthesis". *Evaluation*, *8*(3), 340–358. doi:10.1177/135638902401462448

Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K., (2004). *Realist synthesis: An introduction* (RMP Methods Paper No. 2/2004). Manchester: University of Manchester.

Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2005). Realist review – A new method of systematic review designed for complex policy interventions. *Journal of Health Services Research & Policy*, *10*(1_suppl), 21–34. doi:10.1258/1355819054308530

Penuel, W., Riel, M., Krause, A., & Frank, K. (2009). Analyzing teachers' professional interactions in a school as social capital: A social network approach. *Teachers College Record*, *111*(1), 124–163.

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Malden, MA: Blackwell.

Prestridge, S. (2009). Teachers' talk in professional development activity that supports change in their ICT pedagogical beliefs and practices. *Teacher Development*, *13*(1), 43–55. doi:10.1080/13664530902858493

Raes, E., Kyndt, E., Decuyper, S., Van den Bossche, P., & Dochy, F. (2015). An exploratory study of group development and team learning. *Human Resource Development Quarterly*, *26*(1), 5–30. doi:10.1002/hrdq.21201

*Ritchie, G., & Peters, S. (2001). Using narratives in conferences to improve the CMC learning environment. *Journal of Computer Assisted Learning*, *17*(4), 376–385. doi:10.1046/j.0266-4909.2001.00194.x

*Rock, T. C., & Wilson, C. (2005). Improving teaching through lesson study. *Teacher Education Quarterly*, *32*(1), 77–92.

*Ross, J. A. (1996). The influence of computer communication skills on participation in a computer conferencing course. *Journal of Educational Computing Research*, *15*(1), 37–52. doi:10.2190/6TYW-ADV1-W0LC-N9H0

*Roychoudhury, A., & Roth, W.-M. (1992). Student involvement in learning: Collaboration in science for preservice elementary teachers. *Journal of Science Teacher Education*, *3*(2), 47–52. doi:10.1007/BF02614739

*Ruan, J., & Beach, S. A. (2005). Using online peer dialogue journaling to promote reflection in elementary preservice teachers. *Action in Teacher Education*, *27*(3), 64–75. doi:10.1080/01626620.2005.10463391

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. doi:10.1037/0003-066X.55.1.68

Rycroft-Marlone, J., McCormack, B., Hutchinson, A. M., DeCorby, K., Bucknall, T. K., Kent, B., . . . Wilson, V. (2012). Realist synthesis: Illustrating the method for implementation research. *Implementation Science*, *7*:33. doi:10.1186/1748-5908-7-33

Salas, E., Sims, D. E., & Burke C. S. (2005). Is there a "Big Five" in teamwork? *Small Group Research*, *36*(5), 555–599. doi:10.1177/1046496405277134

Schleicher, A. (2016). *Teaching excellence through professional learning and policy reform: Lessons from around the world*. Paris: OECD. doi:10.1787/9789264252059-en

*Schuck, S. (2003). The use of electronic question and answer forums in mathematics teacher education. *Mathematics Teacher Education and Development*, *5*, 19–31.

*Shin, E.-K., Wilkins, E. A., & Ainsworth, J. (2007). The nature and effectiveness of peer feedback during an early clinical experience in an elementary education program. *Action in Teacher Education*, *28*(4), 40–52. doi:10.1080/01626620.2007.10463428

Showers, B, & Joyce, B. (1996). The evolution of peer coaching. *Educational Leadership*, *53*(6), 12–16.

*Shriki, A., & Movshovitz-Hadar, N. (2011). Nurturing a community of practice through a collaborative design of lesson plans on a wiki system. *Interdisciplinary Journal of E-Learning and Learning Objects*, *7*, 339–357.

*Sluijsmans, D. M. A., Brand-Gruwel, S., Van Merriënboer, J. J. G., & Martens, R. L. (2004). Training teachers in peer-assessment skills: Effects on performance and perceptions. *Innovations in Education and Teaching International*, *41*(1), 59–78. doi:10.1080/1470329032000172720

*So, W. W. (2012). Quality of learning outcomes in an online video-based learning community: Potential and challenges for student teachers. *Asia-Pacific Journal of Teacher Education*, *40*(2), 143–158. doi:10.1080/1359866X.2012.669828

*So, W. W., Hung, V. H., & Yip, W. Y. (2008). The digital video database: A virtual learning community for teacher education. *Australasian Journal of Educational Technology*, *24*(1), 73–90. doi:10.14742/ajet.1231

*Thijs, A., & Van den Berg, E. (2002). Peer coaching as part of a professional development program for science teachers in Botswana. *International Journal of Educational Development*, *22*(1), 55–68. doi:10.1016/S0738-0593(00)00078-X

Thurlings, M., & Den Brok, P. (2017). Learning outcomes of teacher professional development activities: A meta-study. *Educational Review*, *69*(5), 554–576. doi:10.1080/00131911.2017.1281226

Thurlings, M., Evers, A. T., & Vermeulen, M. (2015). Toward a model of explaining teachers' innovative behavior: A literature review. *Review of Educational Research*, *85*(3), 430–471. doi:10.3102/0034654314557949

Topping, K. J. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, *68*(3), 249–276. doi:10.3102/00346543068003249

Van den Bossche, P., Gijselaers, W. H., Segers, M., & Kirschner, P. A. (2006). Social and cognitive factors driving teamwork in collaborative learning environments: Team learning beliefs and behaviors. *Small Group Research*, *37*(5), 490–521. doi:10.1177/1046496406292938

Van Gennip, N. A. E., Segers, M. S. R., & Tillema, H. H. (2009). Peer assessment for learning from a social perspective: The influence of interpersonal variables and structural features. *Educational Research Review*, *4*(1), 41–54. doi:10.1016/j.edurev.2008.11.002

Vangrieken, K., Dochy, F., Raes, E., & Kyndt, E. (2015). Teacher collaboration: A systematic review. *Educational Research Review*, *15*, 17–40. doi:10.1016/j.edurev.2015.04.002

Van Veen, K., Zwart, R., & Meirink, J. (2012). What makes teacher professional development effective? A literature review. In M. Kooy & K. Van Veen (Eds.), *Teacher learning that matters: International perspectives* (pp. 3–21). New York, NY: Routledge.

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. doi:10.1016/j.tate.2007.01.004

*Vincent, J., & Jones, A. (2008). Using models for understanding pedagogical change in a technology environment: A case study of IWB implementation in a secondary school. *Australian Educational Computing*, *23*(2), 32–37. Retrieved from http://acce.edu.au/journal/23/2/using-models-understanding-pedagogical-change-technology-32-environment-case-study-iwb-

*Wade, S. E., & Fauske, J. R. (2004). Dialogue online: Prospective teachers' discourse strategies in computer-mediated discussion. *Reading Research Quarterly*, *39*(2), 134–160. doi:10.1598/RRQ.39.2.1

Wang, T. (2015). Contrived collegiality versus genuine collegiality: Demystifying professional learning communities in Chinese schools. *Compare*, *45*(6), 908–930. doi:10.1080/03057925.2014.952953

*Wilkins, E. A., Shin, E.-K., & Ainsworth, J. (2009). The effects of peer feedback practices with elementary education teacher candidates. *Teacher Education Quarterly*, *36*(2), 79–93.

Wong, G., Greenhalgh, T., Westhorp, G., Buckingham, J., & Pawson, R. (2013). RAMESES publication standards: Realist syntheses. *BMC Medicine*, *11*:21. doi:10.1186/1741-7015-11-21

Wong, G., Westhorp, G., Pawson, R., & Greenhalgh, T. (2013). *Realist synthesis: RAMESES training materials*. Retrieved from http://www.ramesesproject.org/media/Realist_reviews_training_materials.pdf

*Zwart, R. C., Wubbels, Th., Bolhuis, S., & Bergen, Th. C. M. (2008). Teacher learning through reciprocal peer coaching: An analysis of activity sequences. *Teaching and Teacher Education*, *24*(4), 982–1002. doi:10.1016/j.tate.2007.11.003

Appendix 1. Overview of studies

Study	Country	Participants	PTL activity	General type of study	Links to hypotheses	Instrument[a]
Amhag and Jacobson (2009)	Sweden	5 student teachers	collaboration	qualitative	7	Observations (qual)
Amobi (2005)	USA	31 student teachers	coaching	qualitative	5	Reflection report (qual)
					9	Reflection report (qual)
An, Kim, and Kim (2008)	USA	24 in-service teachers	collaboration	qualitative	6	Survey (qual)
					7	Survey (qual)
					9	Survey (qual)
					10	Survey (qual)
Arvaja (2007)	Finland	6 student teachers	collaboration	qualitative	8	Observations (qual)
Baran and Cagiltay (2010)	Turkey	28 student teachers	collaboration	qualitative	5	Reflection report (qual)
					8	Reflection report (qual)
					9	Reflection report (qual)
Biasutti (2011)	Italy	92 in-service teachers	collaboration	mixed methods	1	Survey (quant + qual)
					3	Survey (quant + qual)
					6	Survey (qual)
					9	Survey (quant + qual)
Borko et al. (1997)	USA	14 in-service teachers	collaboration	qualitative	1	Field notes (qual)
					6	Field notes (qual)
					7	Interviews (qual)
					10	Field notes (qual) + interviews (qual)
Bower et al. (2011)	Australia	22 student teachers	coaching	qualitative	1	Survey (quant + qual)
					5	Survey (qual)
					9	Survey (qual)
					10	Survey (qual)
Britton and Anderson (2010)	USA	4 student teachers	coaching	qualitative	10	Interviews (qual) and reflection report (qual)
Buchanan (2011), Buchanan and Stern (2012)	Australia	5 student teachers	assessment	qualitative	1	Survey (qual) and focus groups (qual)
					3	Survey (qual) and focus groups (qual)
					4	Survey (qual) and focus groups (qual)
					6	Survey (qual) and focus groups (qual)
					7	Survey (qual) and focus groups (qual)
					8	Survey (qual) and focus groups (qual)
Burron et al. (1993)	USA	24 student teachers, 27 in control group	collaboration	quantitative	1	Questionnaire (quant)
					2	Observations (quant)
					9	Questionnaire (quant) and observations (quant)
Christ, Arya, and Chiu (2012)	USA	14 in-service teachers	collaboration	mixed methods	1	Observations (qual)
					8	Observations (quantified qual)

Clarke (2009)	United Arab Emirates	unknown number of student teachers[b]	collaboration	qualitative	6	Observations (qual)
					9	Observations (qual)
Donnelly (2005)	USA	18 student teachers	coaching	qualitative	3	Survey (qual)
					8	Survey (qual)
El-Deghaidy and Nouby (2008)	Egypt	14 student teachers; 12 in control	collaboration	mixed methods	8	Survey (quant) and focus groups (qual)
Engelen and Bergen (2002)	Netherlands	37 in-service teachers	coaching	mixed methods	2	Survey (quant) and observations (qual)
					6	Observations (qual)
Estebaranz et al. (1999)	Spain	616 in-service teachers	collaboration	mixed methods	1	Survey (quant and qual)
					7	Survey (quant and qual)
					9	Survey (quant and qual)
Eteläpelto et al. (2005)	Finland	9 student teachers	collaboration	mixed methods	1	Survey (qual)
					6	Survey (qual)
					8	Survey (qual)
					9	Survey (qual)
Forbes (2004)	USA	3 in-service teachers	coaching	qualitative	1	Survey (qual), interviews (qual), and logbooks (qual)
					2	Survey (qual), interviews (qual), and logbooks (qual)
					3	Survey (qual), interviews (qual), and logbooks (qual)
					4	Survey (qual), interviews (qual), and logbooks (qual)
					5	Survey (qual), interviews (qual), and logbooks (qual)
					7	Survey (qual), interviews (qual), and logbooks (qual)
					8	Survey (qual), interviews (qual), and logbooks (qual)
					9	Survey (qual), interviews (qual), and logbooks (qual)
					10	Survey (qual), interviews (qual), and logbooks (qual)
Fransen et al. (2011)	Netherlands	104 student teachers	collaboration	quantitative	6	Focus group (qual)
					9	Survey (quant) and focus group (qual)
Fry and Hin (2006)	Singapore	21 student teachers	coaching	mixed methods	4	Survey (quant)
Goos and Bennison (2008)	Australia	19 student teachers	collaboration	qualitative	6	Observations (qual)
					9	Interviews (qual)

(*Continued*)

Appendix 1. Continued.

Study	Country	Participants	PTL activity	General type of study	Links to hypotheses	Instrument[a]
Grion and Varisco (2007)	Italy	19 student teachers; 18 in-service teachers	collaboration	qualitative	8	Observations (qual)
Hamilton (2013)	USA	28 in-service teachers	coaching	qualitative	1	Survey (quant and qual) and interviews (qual)
					3	Survey (quant and qual) and interviews (qual)
					8	Survey (quant and qual) and interviews (qual)
					9	Survey (quant and qual) and interviews (qual)
Hanewald and Gesthuizen (2009)	Australia	33 in-service teachers	collaboration	qualitative	1	Survey (qual)
					2	Survey (qual)
					7	Survey (qual)
					8	Survey (qual)
					9	Survey (qual)
Harford and MacRuairc (2008)	Ireland	20 student teachers	collaboration	qualitative	1	Focus group (qual)
					2	Focus group (qual)
					3	Focus group (qual)
					5	Focus group (qual)
					7	Focus group (qual)
					8	Focus group (qual)
					10	Focus group (qual)
Harrington and Hathaway (1994) Harrington and Quinn-Leering (1996)	USA	53 student teachers	collaboration	qualitative	3	Observations (qual)
					8	Observations (qual)
Helleve (2007)	Norway	5 student teachers	collaboration	qualitative	6	Observations (qual) and interviews (qual)
					8	Observations (qual) and interviews (qual)
					9	Observations (qual) and interviews (qual)
Hew and Hara (2007)	unclear	20 in-service teachers	collaboration	mixed methods[c]	7	Interviews (qual)
					8	Interviews (qual)
					9	Interviews (qual)
					10	Interviews (qual)
Howard, Barrett, and Frick (2010)	USA	35 student teachers; 37 in control	coaching	qualitative	8	Observations (quantified qual)
Huang et al. (2011)	USA	28 student teachers, 37 in control group	collaboration	qualitative	6	Observations (qual)
					8	Observations (qual) and interviews (qual)
					9	Observations (qual) and interviews (qual)
Hurme et al. (2009)	Finland	6 student teachers	collaboration	qualitative	8	Observations (qual)
					9	Observations (qual) and survey (qual)
Järvenoja and Järvelä (2009)	probably Finland	63 student teachers	collaboration	mixed methods	8	Survey (quant)
					9	Survey (quant)

Jenkins et al. (2005)	USA	37 student teachers	coaching	qualitative	1	Logbook (qual) and peer coaching form (qual)
					2	Logbook (qual) and peer coaching form (qual)
					3	Logbook (qual) and peer coaching form (qual)
Jenkins and Veal (2002)	probably USA	8 student teachers	coaching	qualitative	1	Observations (qual)
					2	Observations (qual)
					3	Observations (qual)
					9	Logbooks (qual)
					10	Observations (qual)
Koc (2011)	Turkey	97 student teachers	collaboration	qualitative	2	Reflection report (qual)
					5	Reflection report (qual)
Lamb et al. (2013)	England	23 student teachers	coaching	qualitative	1	Survey (qual)
					3	Observations (qual), survey (qual), and interviews (qual)
					5	Survey (qual)
					6	Observations (qual), survey (qual), and interviews (qual)
					8	Observations (qual), survey (qual), and interviews (qual)
					9	Observations (qual), survey (qual), and interviews (qual)
Lee and Baek (2012)	probably USA	18 student teachers	collaboration	mixed methods	1	Observations (quantified qual)
					3	Observations (quantified qual)
Lockhorst et al. (2010)	Netherlands	41 student teachers	collaboration	mixed methods	2	Observations (qual) and interviews (qual)
					3	Observations (qual)
					4	Observations (qual)
					10	Observations (qual) and interviews (qual)
Main (2012)	Australia	24 in-service teachers	collaboration	qualitative	8	Observations (qual), interviews (qual), and reflection reports (qual)
					9	Observations (qual), interviews (qual), and reflection reports (qual)
					10	Observations (qual), interviews (qual), and reflection reports (qual)
Mallette et al. (1999)	USA	6 student teachers	coaching	quantitative	2	Observations (quant)
					10	Observations (quant)
Manouchehri (2001)	USA	4 in-service teachers	collaboration	qualitative	1	Observations (qual)
					3	Observations (qual)
					4	Observations (qual) and interviews (qual)
					6	Observations (qual) and interviews (qual)
					8	Observations (qual) and interviews (qual)
					9	Observations (qual, interpretations)

Appendix 1. Continued.

Study	Country	Participants	PTL activity	General type of study	Links to hypotheses	Instrument[a]
Matthew et al. (2009)	USA	37 student teachers	collaboration	qualitative	1	Interviews (qual) and reflection reports (qual)
					2	Interviews (qual) and reflection reports (qual)
					3	Interviews (qual), reflection reports (qual), and observations (qual)
					7	Interviews (qual) and reflection reports (qual)
					8	Interviews (qual) and reflection reports (qual)
					10	Interviews (qual) and reflection reports (qual)
Meirink (2007), Meirink et al. (2010), Meirink et al. (2007), Meirink et al. (2009)	Netherlands	34 in-service teachers	collaboration	mixed methods	1	Logbooks (qual) and observations (qual)
					2	Logbooks (qual) and observations (qual)
					3	Logbooks (qual)
					8	Logbooks (qual), observations (qual), and survey (quant)
					10	Survey (quant)
Nicholas and Ng (2009)	Australia	77 student teachers	collaboration	mixed methods	8	Survey (quant)
					9	Survey (quant) and interviews (qual)
					10	Survey (qual)
Oh and Jonassen (2007)	USA	21 student teachers in first condition, 18 in second condition, 19 in control group	collaboration	mixed methods	8	Observations (qual) and achievement test (qual)
Pachler and Daly (2006)	UK	8 in-service teachers	collaboration	qualitative	8	Interviews (qual)
					9	Interviews (qual)
Prestridge (2009)	Australia	16 in-service teachers	collaboration	qualitative	2	Interviews (qual)
					8	Interviews (qual)
					9	Interviews (qual)
Ritchie and Peters (2001)	New Zealand	57 student teachers	collaboration	qualitative	3	Survey (quant and qual)
					6	Survey (quant and qual)
					8	Survey (quant and qual) and observations (qual)
					9	Survey (quant and qual) and observations (qual)
Rock and Wilson (2005)	USA	6 in-service teachers	collaboration	qualitative	1	Field notes (qual), interviews (qual), and reflection reports (qual)
					2	Interviews (qual) and reflection reports (qual)
					10	Interviews (qual) and reflection reports (qual)
Ross (1996)	USA	15 in-service teachers	collaboration	qualitative	8	Ranking skills (quantified qual) and observations (quantified qual)

					9	Ranking skills (quantified qual) and observations (quantified qual)
Roychoudhury and Roth (1992)	USA	54 student teachers	collaboration	qualitative	1	Observations (qual), field notes (qual), and student reports (qual)
					9	Observations (qual), field notes (qual), and student reports (qual)
Ruan and Beach (2005)	USA	21 student teachers	coaching	qualitative	8	Survey (qual)
					9	Survey (qual)
					10	Survey (qual)
Schuck (2003)	Australia	150 student teachers	collaboration	qualitative	1	Reflection reports (qual)
					2	Reflection reports (qual)
					8	Reflection reports (qual)
					9	Reflection reports (qual)
					10	Reflection reports (qual)
Shin et al. (2007)	USA	56 student teachers	coaching	mixed methods	4	Survey (quant) and interviews (qual)
Wilkins et al. (2009)					7	Survey (quant and qual)
					8	Survey (quant and qual)
					9	Survey (quant and qual) and interviews (qual)
Shriki and Movshovitz-Hadar (2011)	Israel	11 in-service teachers	collaboration	qualitative	2	Observations (qual), interviews (qual), survey (qual), and reflection reports (qual)
					8	Observations (qual), interviews (qual), survey (qual), and reflection reports (qual)
					9	Observations (qual), interviews (qual), survey (qual), and reflection reports (qual)
					10	Observations (qual), interviews (qual), survey (qual), and reflection reports (qual)
Sluijsmans et al. (2004)	Netherlands	50 student teachers; 43 in control group	assessment	quantitative	8	Survey (quant) and interviews (qual)
					10	Interviews (qual)
So (2012)	Hong Kong	25 student teachers	coaching	qualitative	4	Interviews (qual)
					5	Interviews (qual)
					8	Observations (qual)
So et al. (2008)	Hong Kong	3 student teachers	coaching	qualitative	1	Interviews (qual)
					5	Interviews (qual)
					9	Interviews (qual)
					10	Interviews (qual)
Thijs and Van den Berg (2002)	Botswana	120 in-service teachers	coaching	mixed methods	1	Survey (quant) and interview (qual)
					8	Survey (quant) and interview (qual)
					10	Survey (quant) and interview (qual)

Appendix 1. Continued.

Study	Country	Participants	PTL activity	General type of study	Links to hypotheses	Instrument[a]
Vincent and Jones (2008)	Australia	14 in-service teachers	coaching	qualitative	6	Observed (qual; interpretations)
Wade and Fauske (2004)	USA	29 student teachers	collaboration	qualitative	8	Observed (qual)
					9	Observed (qual)
Zwart et al. (2008)	Netherlands	8 in-service teachers	coaching	qualitative	1	Survey (quant), interviews (qual), and logbooks (qual)
					2	Survey (quant), interviews (qual), and logbooks (qual)
					3	Logbooks (qual)
					4	Survey (quant), interviews (qual), and logbooks (qual)

Notes: [a]qual = qualitative; quant = quantitative. [b]Clarke (2009) analysed 750 online messages, but it was not mentioned how many student teachers posted these messages. [c]Hew and Hara (2007) applied a mixed-method design; but only the qualitative measures were relevant for this study.