Teaching and Teacher Education 107 (2021) 103484

Contents lists available at ScienceDirect

Teaching and Teacher Education

journal homepage: www.elsevier.com/locate/tate

Does coaching, mentoring, and supervision matter for pre-service teachers' planning skills and clarity of instruction? A meta-analysis of (quasi-)experimental studies

Sog Yee Mok*, Fritz C. Staub

University of Zurich, Switzerland

HIGHLIGHTS

- Lesson planning and clarity of instruction are pivotal instructional skills for PSTs.
- Coaching, mentoring, & supervision (CMS) had a small effect on instructional skills.
- Cognitive modeling of planning and teaching practices was a significant moderator.
- Experimental CMS studies investigating PSTs' instructional skills are scarce.

ARTICLE INFO

Article history:
Received 16 November 2020
Received in revised form
3 August 2021
Accepted 9 August 2021
Available online 28 August 2021

Keywords:
Coaching
Mentoring
Pre-service teachers
Lesson planning
Clarity of instruction
Meta-analysis

ABSTRACT

During the practicum, pre-service teachers (PSTs) practice instructional skills such as lesson planning and clarity of instruction. Different approaches to assisting PSTs with coaching, mentoring, or supervision have been developed to improve PSTs' instructional skills. We conducted a meta-analysis based on quasi-experimental and experimental studies. The results showed a small and significant overall effect of coaching, mentoring, or supervision on instructional skills ($d = 0.41$). Cooperating teachers' or supervisors' cognitive modeling (i.e., making cognitive processes explicit and demonstrating teaching-related practices) of lesson planning and teaching practices was a significant moderator ($d = 0.90$). Implications for initial teacher education are discussed.

© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Contents

* Corresponding author. Kantonsschulstrasse 3, 8001, Zurich, Switzerland.
E-mail address: sogyee.mok@ife.uzh.ch (S.Y. Mok).

https://doi.org/10.1016/j.tate.2021.103484
0742-051X/© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The practicum is a pivotal phase for pre-service teachers (PSTs) in which they acquire instructional skills and develop professionally (Smith & Lev-Ari, 2005). Instructional skills consist of basic skills like lesson planning and providing clear instructions (i.e., clarity of instruction; Danielson, 2013; van de Grift, 2007). Teachers' instructional practices and skills are associated with cognitive (i.e., professional knowledge) and affective-motivational (i.e., self-efficacy) components of teaching (e.g., Depaepe & König, 2018). As teaching is complex and the attrition rate of novice teachers is high (Hong, 2010), PSTs need to be supported during their initial teacher education and especially during the teaching practicum (Darling-Hammond, Chung, & Frelow, 2002). Various mentoring, coaching, and supervision approaches have been developed to assist and improve PSTs' skill acquisition during their field experience at schools. Narrative reviews have found that learning opportunities within these approaches, such as modeling often provided by cooperating teachers (CTs), have a positive impact on PSTs' professional development during the practicum (Burns, Jacobs, & Yendol-Hoppey, 2016; Hoffman et al., 2015). However, as these various approaches are independently developed by teacher education programs or researchers across the globe, the methods of supporting PSTs are very heterogeneous (e.g., Burns et al., 2016; Lawson et al., 2015). For this reason, the impact of specific mentoring, coaching, and supervision features on PSTs' instructional skills during the practicum are not yet sufficiently understood. The present work addresses this gap by reviewing quasi-experimental and experimental intervention studies and conducting a meta-analysis. The meta-analysis assesses the overall impact of coaching, mentoring, and supervision during the practicum, as well as the effectiveness of different learning opportunities within these approaches on PSTs' instructional skills.

1.1. PSTs' challenges during the practicum

The practicum can positively affect PSTs' feelings of preparedness to teach (Darling-Hammond et al., 2002) and their professional development (Zeichner, 2002). Many studies in the review by Clarke, Triggs, and Nielsen (2014) reported that PSTs evaluate their practicum as one of the most important components of their initial teacher education. During their first practical teaching experiences in schools, however, PSTs may encounter different challenges (Poulou, 2007) that can affect their attainment of skills. For example, at the beginning of their teaching practicum, many tasks and interactions with students are novel to PSTs, which can make lesson planning and the provision of clear instructions demanding for PSTs (Chizhik & Chizhik, 2018a; Danielson, 2013). Moreover, unexperienced PSTs may feel uncertain about their teaching skills, potentially stemming from their low levels of pedagogical content knowledge (Butler & Cuenca, 2012; Costache, Becker, Staub, & Mainhard, 2019). To address these challenges, the support of CTs for PSTs is needed during the practicum (Burns et al., 2016). The term CT is frequently used for teachers who work with PSTs during their practicum (Clarke et al., 2014).

1.2. Mentoring, coaching, and supervision of PSTs

Although many approaches to assisting PSTs during the practicum exist, the approaches explored most are mentoring, coaching, and supervision, which involve a CT or university supervisor (Ambrosetti, Knight, & Dekkers, 2014; Lawson et al., 2015). The terms "mentoring," "coaching," and "supervision" have been used interchangeably in the literature (Ambrosetti et al., 2014). Coaching and mentoring approaches to supporting PSTs differ concerning their focus (Hobson, Ashby, Malderez, & Tomlinson, 2009; Orland-Barak, 2014; Wang & Odell, 2002). In "traditional" mentoring approaches, CTs primarily offer emotional support (e.g., Kemmis, Heikkinen, Fransson, Aspfors, & Edwards-Groves, 2014), and they often discuss general pedagogical topics rather than content-specific issues (Valencia, Martin, Place, & Grossman, 2009). Few mentoring approaches concentrate more on instructional support, such as how to improve lesson planning and teaching quality (Giebelhaus & Bowman, 2002; Gold, 1996). Coaching also concentrates on instructional support, but one difference between coaching and mentoring is that coaching was developed as an on-the-job support for in-service teachers (e.g., Costa & Garmston, 1994; Joyce & Showers, 1981). The effectiveness of coaching for in-service teachers' instructional outcomes has been shown in a recent meta-analysis (Kraft, Blazar, & Hogan, 2018). Less is known about the effectiveness of coaching for PSTs. Some coaching approaches have been adapted to offer PSTs instructional support during the practicum (Becker, Waldis, & Staub, 2019). For instance, Content-Focused Coaching is a coaching approach in which the experienced coach (e.g., the CT) and the less experienced coachee (e.g., the PST) are engaged together in joint lesson planning and the coach provides content-specific feedback (Becker et al., 2019; West & Staub, 2003). The terms "supervision" or "clinical supervision" are used to describe the collaboration in a hierarchical relationship between CTs or university teachers and PSTs during their first field placement (Lawson et al., 2015; Metcalf, 1992). Most of the studies that have investigated the benefits of mentoring, coaching, or supervision on PSTs' instructional skills are qualitative (Allen, 2011; Lawson et al., 2015); quantitative studies are scarce (Burns et al., 2016). Narrative reviews based on qualitative and quantitative studies describe a number of factors such as assistance, individual support, and curriculum practices of supervisors or CTs during the practicum to foster PSTs instructional skills (Burns et al., 2016; Hoffman et al., 2015). Increasing the number of randomized or quasi-experimental studies with a control group design has been recommended to allow for more evidence-based conclusions (Cook, Campbell, & Shadish, 2002; Slavin, 2008) about effective mentoring approaches (Smith & Ingersoll, 2004). Indeed, there is a lack of research assessing the effectiveness of coaching, mentoring, and supervision for PSTs' instructional skills during the practicum based on experimental and quasi-experimental studies.

1.3. Development of PSTs' skills for lesson planning and clarity of instruction during the practicum

The objectives of instructional skills training during the practicum can be very heterogeneous across and within countries (Arnold, Gröschner, & Hascher, 2014; Ronfeldt & Reininger, 2012; Ulrich et al., 2020). Although the objectives vary across countries and federal states (Ronfeldt & Reininger, 2012; Ulrich et al., 2020), two common aspects can be highlighted: planning a lesson as part of the university coursework and planning and conducting a lesson as a concrete action in the classroom during the practicum (Danielson, 2013; König et al., 2017; Pecheone & Chung, 2006).

In general, effective teaching is widely conceptualized to encompass teaching behavior or teaching patterns that positively affect the processes and outcomes of student learning (Kyriakides, Christoforou, & Charalambous, 2013; Seidel & Shavelson, 2007). For example, a meta-analysis by Kyriakides et al. (2013) suggested that effective teaching behavior consisted of different aspects, such as orientation (i.e., providing objectives for specific lessons), structuring (i.e., providing learning materials and arranging the material in meaningful order), management of time, and questioning. For unexperienced PSTs achieving all aspects of effective teaching during their first practical experiences in schools is demanding. Research has proposed that PSTs' teaching skill development should start with the basic tasks of lesson planning and providing clear instructions (i.e., including objectives and clear structures) before working on more advanced tasks such as making use of higher-order thinking techniques for students' cognitive activation (Fuller & Bown, 1975; Tas, Houtveen, van the Grift, & Willemsen, 2018). Stimulating cognitive activation requires PSTs to have the capacity to focus on students' individual learning needs (Baumert et al., 2010; Tas et al., 2018).

Lesson planning, which is a "form of constructing mental representations of possible classroom settings" (Seidel, Blomberg, & Renkl, 2013, p. 56), is the first step of a full cycle of teaching for PSTs (Seidel et al., 2013). In various European initial teacher education programs, lesson planning is a crucial task (European Commission, 2013). Through lesson planning, PSTs learn to structure and prepare their lessons (European Commission, 2013; Flores, 2016). Similarly, in the U.S., PSTs' lesson planning skills are examined and evaluated at the end of the initial teacher education, such as by the PACT (Performance Assessment for California Teachers) assessment (Pecheone & Chung, 2006). Successful lesson planning reflects procedures that are aligned with students' domain-specific knowledge (König, Bremerich-Vos, Buchholtz, Fladung, & Glutsch, 2020). However, PSTs often fail to consider the individual learning needs of their students when planning a lesson (Chizhik & Chizhik, 2018a).

Clarity of instruction is a key skill that is highlighted by all three of the well-established instruments for assessing teaching and instruction (Maulana, Helms-Lorenz, & van de Grift, 2015): "The Classroom Assessment Scoring System" (CLASS; Pianta, La Paro, & Hamre, 2008; Pianta & Hamre, 2009), "The Framework for Teaching Evaluation Instrument" (Danielson, 2013), and "The International Comparative Analysis of Learning and Teaching instrument" (ICALT, van de Grift, 2014). Clarity of instruction entails providing clear learning goals and instructions (Danielson, 2013; Metcalf, 1992). In the U.S., being able to provide clear instruction is additionally essential for PSTs because it is tested and evaluated at the end of initial teacher education (Pecheone & Chung, 2006).

To sum up, as previous research has shown that lesson planning and providing clear instructions are pivotal learning outcomes for PSTs during the practicum (e.g., Pecheone & Chung, 2006), we focus on both lesson planning and clarity of instruction as the key instructional skills which PSTs should accomplish during initial teacher education. Thus, CTs need to assist PSTs in their development of both lesson planning and clarity of instruction.

1.4. CTs' support for opportunities for learning

A vast amount of research has explored the collaboration between CTs and PSTs for the development of teaching behavior (Clarke et al., 2014; Le Cornu & Ewing, 2008; Mena, Hennissen, & Loughran, 2017; Orland-Barak, 2014). The literature has identified different roles that CTs can play to assist in PSTs' learning during the practicum. In their review, Clarke et al. (2014) differentiate three basic conceptions of the CT's role in teacher education: In the first conception, the CT is seen as a classroom placeholder who participates to a minimal degree in the lessons only (Clarke et al., 2014). In the second conception, the CT takes on the role of a supervisor during the practicum, overseeing the work of the PST. The CT has to observe, record, and report the successful use of knowledge in the practice of teaching in classrooms (Borko & Mayfield, 1995; Clarke et al., 2014). The level of participation is higher for CTs as supervisors than as classroom placeholders. Interactions between the CT and PST are rather unidirectional and driven by the CT. In the third conception, the CT is seen as a teacher educator, with a high level of participation. As teacher educators, CTs act as coaches who work closely with learners and provide guidance to foster the development of the PST's teaching repertoire through collaboration with the PST. Thus, these conceptions of CTs' roles are linked to different activities to assist PSTs' learning, such as providing feedback, modeling of practices, and supporting PSTs' reflections (Clarke et al., 2014). These activities and CTs' level of participation are affected by their felt responsibility for students' learning in the classroom and the desire to assist and support unexperienced PSTs in developing their skills. Drawing on the conceptions by Clarke et al. (2014), we assume that PSTs' learning experiences of coaching, mentoring, and supervision are related to the level of participation which affects the instructional support of a CT or supervisor.

1.5. Learning opportunities as potential moderators

Various learning opportunities provided by a CT or supervisor have been shown to improve PSTs' teaching skills (Burns et al., 2016; Fuentes-Abeledo, González-Sanmamed, Muñoz-Carril, & Veiga-Rio, 2020; McDonald, Kazemi, & Kavanagh, 2013). In recent years, researchers have additionally begun to incorporate the use of digital tools in coaching, mentoring, and supervision to provide more flexible feedback (Burns et al., 2016) and to draw on student peers as an additional resource for feedback and fostering reflection (Lu, 2010). However, no meta-analytic work has investigated types of learning opportunities in coaching, mentoring, and supervision as moderators that potentially affect the strength of the overall effect on instructional skills (i.e., lesson planning and clarity of instruction). Thus, the following section identifies learning opportunities that could serve as potential moderators of different approaches to assisting PSTs in the practicum on the overall effect on instructional skills.

Learning opportunities in approaches to assisting PSTs. Among the different approaches to coaching, mentoring, and clinical supervision there is broad agreement in distinguishing three phases in a cycle of assisting novice teachers: in pre-lesson conferences, during the enactment of lessons, and in post-lesson conferences (e.g., Goldhammer, Anderson, & Krajewski, 1969; Hoffman et al., 2015; Staub, 2004). These phases vary with respect to the learning opportunities they provide for PSTs (McDonald et al., 2013). Nevertheless, due to the different needs of PSTs, the extent to which these various learning opportunities can help PSTs during the practicum is still an open question.

Learning opportunities during pre-lesson conferences. Pre-lesson conferences aim to provide opportunities for conversations between the CT and PST in *pre-lesson dialogues* to help unexperienced PSTs with lesson planning (Futter, 2017). When these dialogues concern mutual lesson planning, they are also known as co-planning (Smith, 2005). Research suggests that pre-lesson conferences enable CTs and PSTs to plan lessons together through co-constructive dialogues in which both coach and coachee accept *joint responsibility for the lesson* (Becker et al., 2019; West & Staub, 2003). Based on this mutual responsibility, PSTs and CTs both contribute and discuss their ideas on how the lesson should be designed to foster student learning (Staub, 2004).

One of the most prominent theoretical models in the mentoring and coaching literature is the seminal cognitive apprenticeship model by Collins, Brown, and Holum (1991), in which *modeling* is a core element. In modeling, a less experienced coachee learns skills from an expert that are not fully observable, such as cognitive processes reflecting procedural and conceptual knowledge (Collins, Brown, & Newman, 1989). By observing and reflecting on the practice and cognitive processes modeled by the experienced coach, the coachee can learn new a practice and build a cognitive model of the practice (Collins et al., 1989).

The cognitive apprenticeship model framework was initially developed to describe school students' learning processes guided by their teachers (Brown, Collins, & Duguid, 1989). The cognitive apprenticeship model (Brown et al., 1989) was adapted to the teacher education context to advance teacher education programs (Dennen & Burner, 2008) by improving lesson planning skills (Liu, 2005) and teaching behavior (Jager, Reezigt, & Creemers, 2002). The cognitive apprenticeship methods such as modeling were, therefore, modified to the requirements of practicum and core practices of PST education (McDonald et al., 2013). Based on previous research, we assume modeling to be one of the key mentoring and coaching strategies by which CTs support PSTs in developing their teaching competencies during field experiences (Clarke et al., 2014; Ellis, Alonzo, & Nguyen, 2020; McDonald et al., 2013). We call the modeling introduced by Collins et al. (1991) *cognitive modeling* in our work because modeling in teacher education consists of two processes: First, making cognitive processes explicit (e.g., by discussing or deliberating on why the CT or supervisor has chosen a certain task in a lesson plan or in the lesson) and second, demonstrating effective teaching-related practices based on the previously explained cognitive processes. Based on these two processes, two forms of cognitive modeling are commonly used in the initial teacher education (Ellis et al., 2020; Smith, 2005). The first form refers to lesson planning: CTs or supervisors demonstrate how to effectively plan a lesson while making the cognitive processes underlying the decisions to a lesson plan explicit. The second form of cognitive modeling refers to the demonstration of teaching practices enacted during lessons and the reasoning behind the decisions for the selected teaching practices. The demonstration of teaching practices during lessons can be observed by the PSTs (McDonald et al., 2013), or gained from viewing videotaped lessons after the lesson (Grosser-Clarkson & Neel, 2019; Liu, 2005). The reasoning underlying the decisions to a lesson plan or enacted teaching practices can be deliberated in pre-conferences (Smith, 2005) or while analyzing the videotaped lesson in post-conferences (Ellis et al., 2020; Liu, 2005).

Learning opportunities during the lesson. Research has shown that the opportunity to observe the CTs' practice in their classrooms is a prerequisite for PSTs' learning during the practicum (Grosser-Clarkson & Neel, 2019). For example, PSTs during their first field experience can learn effective teaching behavior from their experienced CTs by *personal observation* of the CT's expert teaching behavior in real classrooms, which can also stimulate PSTs' reflection of their teaching practices (Ronfeldt, Brockman, & Campbell, 2018). Thus, we additionally focused on the impact of PSTs' opportunity to observe the CTs' teaching practices on instructional skills.

Learning opportunities during post-lesson conferences. *Reflection* is an essential focus of teacher education (Bengtsson, 1995; Marcos, Miguel, & Tillema, 2009) because it can improve the quality of teaching (Schön, 1983) and is pivotal for professional development (Darling-Hammond, 2014). In post-lesson conferences, CTs often focus on the evaluation of the past lesson, rather than giving PSTs the opportunity to reflect on their own practice in-depth for a next lesson (Douglas, 2011; Hoffman et al., 2015). Hence, post-lesson conferences might be less effective for PSTs' learning when self-reflection opportunities are not provided (Gibbons & Cobb, 2017). Thus, opportunities for reflection are important for PSTs' instructional skill development. Feedback from CTs on PSTs' teaching and learning processes are also necessary to ensure the improvement of teaching skills (Lee & Wu, 2006). *Feedback* consists of information about the PSTs' skills during the learning process and how they can develop these skills in future practice (Hattie & Timperley, 2007; Thurlings, Vermeulen, Bastiaens, & Stijnen, 2013). Recent research suggested that more structured observation based on observation tools can be helpful to provide more systematic feedback on PSTs' teaching practices (Tas et al., 2018). The feedback of CTs can be used by PSTs after the lesson to improve future lesson plans and lessons.

The supplementary role of digital tools and peers in coaching, mentoring, and supervision. The use of digital tools and peer support during the practicum have emerged in the last years to make the collaboration between CTs and PSTs more efficient and to allow for more flexibility in supporting PSTs' learning (Barnett, Keating, Harwood, & Saam, 2002; Burns et al., 2016). An increasing number of studies have investigated how PSTs advance their instructional skills through coaching, mentoring, and supervision that integrates *digital tools* to provide real-time feedback in coaching and mentoring, like online video-conferences, video-based feedback, and online discussion forums (Burns et al., 2016; Kopcha & Alger, 2014; Rock et al., 2012; Weber, Gold, Prilop, & Kleinknecht, 2018). These technology-based elements can offer additional support and different sources of feedback for PSTs' learning during the practicum (Burns et al., 2016).

Supplementary forms of assistance to coaching, mentoring, and supervision, such as *peer coaching*, have emerged to reduce the workload of CTs and supervisors during the practicum (Kreis, 2019; Lu, 2010). In peer coaching the CT or supervisor has a less active role. During peer coaching, students at a similar competence level exchange and discuss their ideas for planning a lesson in a pre-lesson conference or what they observed during the lesson in a post-lesson conference (Lu, 2010). In most studies the CT or supervisor is absent during such peer conferences but available to answer PSTs' unsolved questions (e.g., Nguyen & Baldauf, 2010), but in some studies the CT's role is not explicitly described (e.g., Bowman & McCormick, 2000). Nevertheless, all of these studies explicitly or implicitly imply a minimal involvement of the CT or supervisor during the practicum. Studies have shown that peers can provide support during their collaboration with regard to teaching-related activities (Baeten & Simons, 2016; Kreis, 2019). A review by Lu (2010) showed that many qualitative studies support the effectiveness of peer coaching for PSTs' instructional skills, but quantitative results are scarce. In sum, learning opportunities, digital tools, and the additional support of peer coaches are promising areas of research, but the extent to which they moderate the overall effect of coaching, mentoring, and supervision on PSTs' instructional skills has not yet been explored.

1.6. The present study

The study has two aims. The first aim of the present meta-analysis is to investigate the effectiveness of coaching, mentoring, and supervision in the teaching practicum on PSTs' instructional skills (i.e., lesson planning and clarity of instruction). The second aim is to evaluate which moderator variables affect the overall effectiveness of coaching, mentoring, and supervision. Accordingly, the meta-analysis was guided by two research questions (RQs):

RQ 1) How strong is the overall effect of coaching, mentoring, and supervision interventions on PSTs' instructional skills during the practicum?

RQ 2) What moderator variables affect the overall effect of coaching, mentoring, and supervision interventions on PSTs' instructional skills during the practicum?

2. Method

2.1. Literature search

We conducted a systematic literature search in the databases Educational Resources Information Center (ERIC), PsychINFO, Web of Science, and Google Scholar. For the systematic literature search, we used different combinations of search terms to identify intervention studies. To identify different approaches to assisting and fostering PSTs in the practicum, we used the terms coaching, mentoring, supervision, training, or intervention. To detect the target group of PSTs, we used the terms student teachers, teacher candidates, and PSTs. We applied the terms instructional skills, instructions, and teaching skills to achieve a large initial sample and to search for the relevant outcomes within the coaching, mentoring, and supervision studies for PSTs. The search was limited to the abstract and title of peer-reviewed articles in English and German language. We included German articles to maximize the number of eligible studies. We additionally checked the abstracts of potential articles published in significant teacher education journals such as *Teaching and Teacher Education*. The reference list of previous reviews (i.e., Burns et al., 2016; Hoffman et al., 2015) were also examined for potential articles. The review by Burns et al. (2016) included studies published between 2001 and 2013 that explored the PSTs' supervision within field experiences. Burns et al. (2016) narratively summarized 69 (mostly qualitative) studies in their review. Hoffman et al. (2015) focused in their review on coaching interactions between CTs and PSTs in teacher education programs and how these interactions can affect learning. Hoffman et al. (2015) narratively reviewed 46 studies in which 76 % used a qualitative design, 7 % used mixed methods, 15 % used a quantitative descriptive design, and only 2 % used a quantitative experimental study design.

2.2. Inclusion and exclusion criteria

To be included in the analyses, studies had to investigate a coaching, mentoring, or supervision intervention for PSTs aimed at improving PSTs' lesson planning and/or clarity of instruction. Studies also had to use a quasi-experimental or experimental pre- and post-test design or a post-test design with a control group. Studies were excluded from the analyses if they a) concentrated on sports education or trained PSTs for students with special needs or low-income settings, as the teaching content might deviate strongly from the other studies, or b) did not provide sufficient data to calculate relevant effect sizes (despite the authors being contacted and asked to share the data for the meta-analysis).

2.3. Coding

We identified 386 potential studies from the literature search. In a first step, the first author screened all abstracts based on the inclusion and exclusion criteria. This screening resulted in 39 studies. 20 % of abstracts were independently coded by a second rater. The interrater-reliability was good (Cohen's $\kappa = 0.89$). The few disagreements were discussed and the raters agreed on a solution. In a second step, the full texts of the remaining 39 studies were examined, resulting in 14 eligible studies for our meta-analysis. Reasons studies were excluded from the meta-analysis at this point included measuring frequencies only at the end of PSTs' program (instead of means and standardizations) or not applying an experimental design with a control group. All study characteristics (e.g., type of intervention, design, outcomes, place of intervention) and potential moderators (dummy-coded) were coded by the first author. Approximately 20 % of all study characteristics and moderators were additionally coded by a trained rater. The interrater agreement reached 88.89 %, indicating few disagreements. The disagreements between raters were discussed and resolved.

2.4. Calculation of overall effect sizes

While we can calculate single effect sizes from one specific context in single studies, we gain more robust evidence for the effectiveness of the interventions in these studies by pooling the effects to calculate an overall effect across studies (Borenstein et al., 2021). Thus, we explored the effectiveness of coaching, mentoring, and supervision to improve PSTs' instructional skills by calculating the overall (pooled) effect size across multiple approaches to coaching, mentoring, and supervision. The effect size calculations and subgroup analysis (i.e., moderator analysis) were conducted in *R* (R Core Team, 2015) with the R-packages *metafor* (Viechtbauer, 2010) and *esc* (Lüdecke, 2018).

As we were interested in quasi-experimental or experimental studies with either a pre-posttest-control group design or a posttest-control group design, we calculated the effect sizes from each individual study before we calculated the mean overall effect. As most group sizes within studies were balanced, we used Cohen's *d* as the effect size measure. If studies reported insufficient statistical metrics to calculate effect sizes, we contacted the authors of the original studies and asked for the missing metrics; if authors did not provide the necessary information, we omitted these studies from the final sample. In one study (Kopcha & Alger, 2011) the missing *SD*s were replaced with *SD*s reported in a follow-up study from the same authors (Kopcha & Alger, 2014). For studies with a pre-posttest-control group design, we calculated the effect sizes following the recommendations by Morris (2008). For studies with a posttest-control group design, we calculated the effect sizes if the means and standard deviations of the experimental and control group or *F*-statistics and the *p*-value of an ANOVA were reported. We then proceeded with the calculation of the overall effect across all studies using either a pre-posttest-control group design or posttest-control group design for the meta-analysis. We chose a random-effects model to calculate the mean overall effects because we assumed that the true effect size could vary across studies (Borenstein et al., 2021; Lipsey & Wilson, 2001). With regard to the calculations of the pooled effect of instructional skills, we included studies that either reported measures of planning skills or clarity of instruction. If studies explored both outcomes, we used an average measure of both outcomes to avoid the dependence of measures within one study (Hunter & Schmidt, 2004). If studies investigated two intervention groups, we also used an average measure of the outcomes from both interventions groups and compared it to the control group. To gain insight in the different components of instructional skills, we further calculated separate overall effects for the planning skills and clarity of instruction sub-samples.

We conducted an influence analysis to avoid the risk of bias from extreme outliers that do not fit the meta-analytic model of the sample (Viechtbauer, 2010). If these studies with extreme outliers were included in the sample, the overall effect across studies would be biased towards an inflated effect (Borenstein et al., 2021). To detect potential outliers, we conducted influence analyses on the full sample (Viechtbauer & Cheung, 2010). To do so, we calculated the standardized deleted residuals (SDRs), which indicate the deviation of the correlation of a study from the average correlation of the other studies. If studies reached SDRs values below −1.96 or above 1.96, they were considered outliers. For the additional overall effects, we omitted the influence analyses if the sub-samples were smaller than ten studies because the analysis is not robust in such small samples (Viechtbauer & Cheung, 2010).

2.5. Moderator analyses

After calculating the overall effect, we conducted a heterogeneity test to explore if there was a substantial variability in true effect sizes between studies. A moderate (i.e., I^2 of 50–79%) or high variability (i.e., $I^2 > 80\%$) of effect sizes is a prerequisite for conducting a moderator analysis to explain the variability of effect sizes (Higgins & Thompson, 2002). We used mixed-effects models for the moderator analyses (Viechtbauer, 2010) of all potential moderators (i.e., dialogues during pre-lesson conference, joint responsibility between coach and coachee, cognitive modeling, personal observation, opportunities for reflection for PSTs, provide structured feedback, digital tools, and peer coaching). In all mixed-effect models, we included only single moderators due to the small sample size. The moderator analysis followed two steps (Viechtbauer, 2010). First, an omnibus test indicates if the moderator has a significant influence on the outcome (i.e., PSTs' instructional skills) and if the categorical levels of a moderator differ significantly. Second, the effect size and significance are calculated post hoc for each moderator level.

3. Results

3.1. Descriptive statistics

The descriptive statistics are displayed in Table 1. The subject domains in the studies varied from language ($k = 5$), mathematics ($k = 4$), biology ($k = 1$), and to multiple subject domains within studies ($k = 2$). Four studies did not explicitly report the subject domain. With regard to the school level, eight studies explored PSTs of secondary school students and eight studies investigated PSTs of elementary school students. It should be noted that two studies included PSTs from both secondary and elementary level. The majority of studies ($k = 7$) focused on coaching performed by CTs; others applied peer coaching approaches ($k = 4$). Fewer scholars explored mentoring ($k = 2$) or supervision ($k = 2$) in their studies. It should be noted that some studies combined different approaches (e.g., supervision and other learning community approaches such as "lesson study"). Only five studies investigated digital tools such as online discussion boards (e.g., Kopcha & Alger, 2014) or e-conferences (e.g., Liu, 2005) within their assisting approach. Many studies investigated either lesson planning ($k = 8$) or clarity of instruction ($k = 11$). Only five studies explored both outcomes (e.g., Chizhik, Chizhik, Close, & Gallego, 2017; cf. Table 1).

3.2. Overall effects

Fourteen studies focusing on instructional skills (i.e., including planning skills, instructional skills, or both) were initially included in the meta-analysis. The meta-analysis showed an overall effect for all 14 studies with a high variability of effect sizes between studies ($d = 0.66$, $SE = 0.19$, $z = 3.43$, $p < .001$, $CI[0.28; 1.03]$; see Online Appendix Figure A1). We investigated potential outliers by conducting influence analyses (Viechtbauer & Cheung, 2010). Following the influence analyses, we detected two extreme outliers which both exceeded the SDRs limit (Goker, 2006, $z = 2.18$; Bowman & McCormick, 2000, $z = 2.65$, see Online Appendix Figure A2). Consequently, we excluded both studies with very large effects ($ds > 2.0$) from the final sample and re-ran the meta-analysis. The final overall effect of coaching, mentoring, and supervision interventions on instructional skills was significant and small ($d = 0.41$, $SE = 0.13$, $z = 3.22$, $p = .001$, $CI[0.16; 0.66]$, $k = 12$; RQ 1: How strong is the overall effect of coaching, mentoring, and supervision interventions on PSTs' instructional skills during the practicum?). The overall effect on instructional skills including the confidence intervals are displayed in a forest plot (Fig. 1).

As instructional skills consisted of two components (i.e., planning and clarity of instruction), we additionally calculated the overall effects for both components separately in two sub-analyses. The overall effect for PSTs' *planning skills* was small but not significant ($d = 0.28$, $SE = 0.17$, $z = 1.64$, $p = .102$, $CI[-0.06; 0.62]$, $k = 8$; Fig. 2). In contrast, the overall effect of studies investigating *clarity of instruction* was significant and small-to-medium ($d = 0.49$, $SE = 0.09$, $z = 5.58$, $p < .001$, $CI[0.32; 0.66]$, $k = 9$; Fig. 3).

3.3. Moderator effects

The test of heterogeneity regarding the instructional skills was significant ($Q = 35.84$, $df = 11$, $p < .001$, $I^2 = 67.61$) and the I^2 of 67.61 % indicated a moderate level of variability of effect sizes across studies (Higgins & Thompson, 2002). Thus, the moderator analysis on instructional skills with the full sample could be conducted. The results of the moderator analysis on the instructional skills effects showed that cognitive modeling was the only significant moderator (RQ 2: What moderator variables affect the overall effect of coaching, mentoring, and supervision interventions on PSTs' instructional skills during the practicum?): PSTs showed better instructional skills when their CTs or supervisors made use of cognitive modeling ($d = 0.90$, $p < .001$) compared to when CTs or supervisors did not make use of cognitive modeling ($d = 0.28$, $p = .03$; $Q = 4.53$, $df = 1$, $p = .033$, Table 2). Cognitive modeling showed a large effect size of $d = 0.90$. All other moderators (i.e., learning opportunities) were not significant. The moderators associated with each of the three phases during coaching, mentoring, or supervision (i.e., pre-lesson, during the lesson, or after the lesson) are displayed in Table 2. The moderators digital tools and peer coaching which cannot be assigned to specific phases are displayed in Table 3.

Table 1
Descriptive statistics.

Study	Country	Type of support/Use of digital elements	Theoretical model of coaching, mentoring, or supervision	Study design	Dependent variables	Sample size (PSTs)	Domain/school level	Length of coaching, mentoring, or supervision for PSTs
Becker, Waldis, & Staub (2019)	Switzerland	Coaching	Adapted Content-Focused Coaching by West and Staub (2003)	Ex, Pr-Po	Planning & Instruction	59[a]	Mathematics/ secondary	3 weeks
Bowman & McCormick (2000)	USA	Peer coaching	Based on Joyce & Showers (1981)	Ex, Pr-Po	Instruction	32	English/elementary school	6 weeks
Chizhik & Chizhik (2018b)	USA	Supervision/ digital	Based on cultural historical activity theory by Engestrom (1999)	Ex, Po	Planning & Instruction	54	Multiple- subjects/ elementary	52 weeks (occasional support)
Chizhik et al. (2017)	USA	Supervision & others/digital	Lesson study by e.g., Stigler & Hiebert (1999)	Q-ex, Po	Planning & instruction	60	Multiple- subjects/ elementary	52 weeks (occasional meetings)
Giebelhaus & Bowman (2002)	USA	Mentoring	Praxis III/Pathwise model, e.g., Jones (1992)	Q-ex, Po	Instruction	29	English, mathematics, social studies/elementary & secondary	10 weeks
Goker (2006)	Northern Cyrus	Peer coaching	Based on Joyce & Showers (1981)	Ex, Pr-Po	Instruction	32	English/elementary school & secondary school	6 weeks
Kopcha & Alger (2011)	USA	Coaching/ digital	Adapted cognitive apprenticeship by Collins et al. (1991)	Q-ex, Po	Planning & instruction	60	NA/secondary	26 weeks (over 1 semester)
Kopcha & Alger (2014)	USA	Coaching/ digital	Cognitive apprenticeship by Collins et al. (1991)	Q-ex, Po	Planning & instruction	54	NA/secondary	10–12 observations over 52 weeks
Kreis & Staub (2011)	Switzerland	Coaching	Content-Focused Coaching by West and Staub (2003)	Q-ex, Po	Instruction	32	Mathematics/ elementary	7 weeks
Liu (2005)	Taiwan	Coaching/ digital	Cognitive apprenticeship by Collins et al. (1991)	Ex, Pr-Po	Planning	24	NA/elementary	7 weeks
Nguyen & Baldauf, 2010	Vietnam	Peer coaching & mentoring	Drawn from effectiveness studies of the peer mentoring process (e.g., Bryant & Terborg (2008))	Q-ex, Pr-Po	Instruction	65	English/secondary”	6 weeks
Smit, Rietz, & Kreis, 2018 [b]	Switzerland & Germany	Peer coaching	Adapted Content-Focused Coaching by West and Staub (2003)	Q-ex, Po	Planning	118	Biology/secondary	4 weeks
Tas, Houtveena, van de Grift, & Willemsen (2018)	Netherlands	Coaching	Based on different theories on professional development of (student) teachers: concern theory by Fuller & Bown (1975); research on effective teaching behavior: e.g., (Hanushek (2011); Scheerens (2015))	Q-ex, Pr-Po	Instruction	198	Language or mathematics/ Kindergarten & elementary	52 weeks
Veeman, Denessen, Gerrits, & Kenter (2001)	Netherlands	Coaching	Based on Goldhammer (1969), Cogan (1973), Goldhammer, Anderson, & Krajewski (1993), Joyce & Showers (1981) and Costa and Garmston (1994)	Q-ex, Pr-Po	Planning[c]	34	NA/secondary	14 weeks

Note. CTs = Cooperating teachers; PSTs = Pre-service teachers; NA = not available; Q-ex = Quasi-experimental design, Ex = Experimental design; Po = Post design, Pr-Po = Pre-post design.

[a] We combined the effects of the three experimental groups used in Becker et al. (2019). Thus, the calculated final sample size included in our meta-analysis was reduced to 32.

[b] Smit et al. reported that the distribution of female and older more experienced university students was higher in the control group (after we requested relevant unpublished data via e-mail). Thus, we were advised to control for gender and semester in the analysis.

[c] Lesson planning of PSTs was measured indirectly by the sub-scale making concrete improvement plans for CTs in Veenman et al. (2001). As the coaching was based on (co-constructive) principles such mutual trust, enhancing autonomy, and improvement of instructional practice by providing feedback and stimulating PSTs to be more reflective in lesson-conferences, it can be assumed that such co-constructive planning of CTs can affect PSTs' planning skills.

4. Discussion

The present meta-analysis summarizes findings from quasi-experimental and experimental studies examining the effects of coaching, mentoring, and supervision designed to improve PST's instructional skills during their teaching practicum. Regarding our first research question (i.e., how strong is the overall effect of coaching, mentoring, and supervision interventions on PSTs' instructional skills during the practicum?), we showed a significant and small overall effect ($d = 0.41$, $k = 12$) for the included intervention studies on PSTs' instructional skills (i.e., lesson planning skills and/or clarity of instruction) even after removing two outliers from the final sample. When analyzing the overall effects for lesson planning skills and clarity of instruction separately, we found a non-significant small overall effect for planning skills ($d = 0.28$) and a significant and small-to-medium overall effect for clarity of instruction ($d = 0.49$). With regard to our second research question (i.e., what moderator variables affect the overall effect of coaching, mentoring, and supervision interventions on PSTs' instructional skills during the practicum?), cognitive modeling was the only significant moderator of the effect on instructional skills. Cognitive modeling had a large effect on PSTs' instructional skills ($d = 0.90$). Thus, cognitive modeling is a key element in assisting PSTs in the practicum that can amplify the effectiveness of coaching, mentoring, and supervision for PSTs' instructional skills. To the best of our knowledge, our meta-analysis is the first to investigate the effectiveness of coaching, mentoring, and supervision during the practicum to promote PSTs' instructional skills.

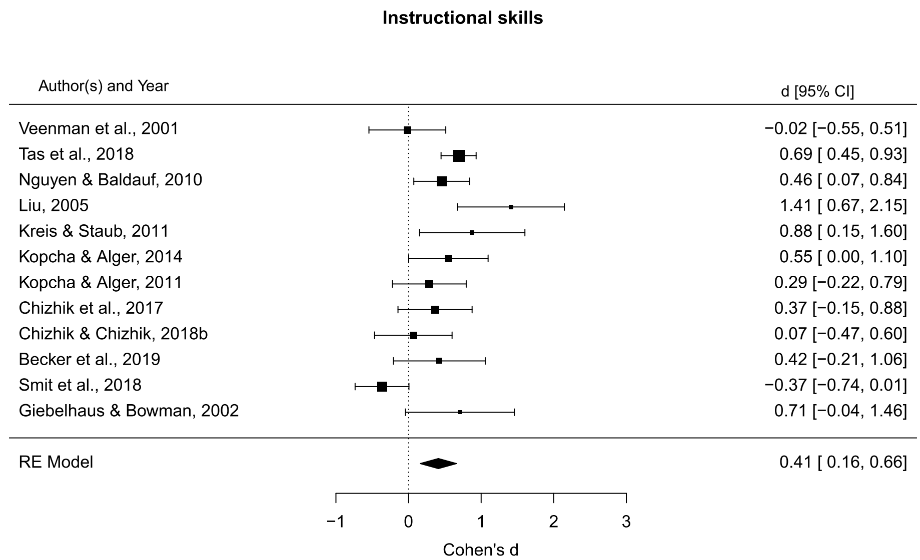

Fig. 1. Overall effect for instructional skills (final sample: $k = 12$ studies).

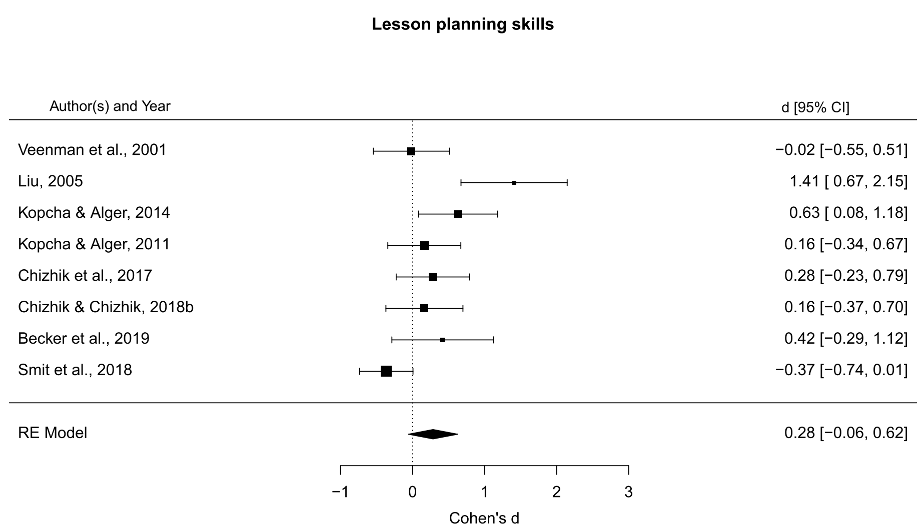

Fig. 2. Overall effect for lesson planning skills ($k = 8$ studies).

4.1. *Relevance of quantitative evidence for the effectiveness of coaching, mentoring, and supervision interventions on PSTs' instructional skills*

Our findings are in line with the results of earlier narrative research syntheses. Previous reviews have proposed benefits of collaboration between CTs and PSTs in initial teacher education (Burns et al., 2016; Hoffman et al., 2015; Lawson et al., 2015), which has been supported by the positive effects we found on instructional skills. In contrast to our meta-analysis, these reviews based their findings on mostly qualitative studies or studies with survey-based designs without a control group. A recent meta-analysis for in-service teachers found a large averaged coaching effect of 0.49 standard deviations on in-service teachers' instructional skills (Kraft et al., 2018). The effect size we observed for the overall effect on instructional skills (and clarity of instruction in the sub-analysis) for PSTs is less strong (i.e., small) in comparison to the coaching effect for in-service teachers (i.e., large; Kraft et al., 2018). Our smaller effect size on instructional skills might be explained by PSTs' lack of prior teaching experience and the relatively short period of the teaching practicum in schools, which may make it more difficult for PSTs to successfully advance their planning skills and particularly clarity of instruction using coaching as compared to experienced in-service teachers. The present meta-analysis also adds to existing teacher education research by investigating the effectiveness of coaching, mentoring, and supervision studies for PSTs' instructional skills based on quantitative studies using a quasi-experimental or experimental control-group design. Our findings provided evidence that coaching, mentoring, and supervision with trained CTs are more effective than traditional approaches without trained CTs.

Importantly, many unexperienced PSTs (and beginning

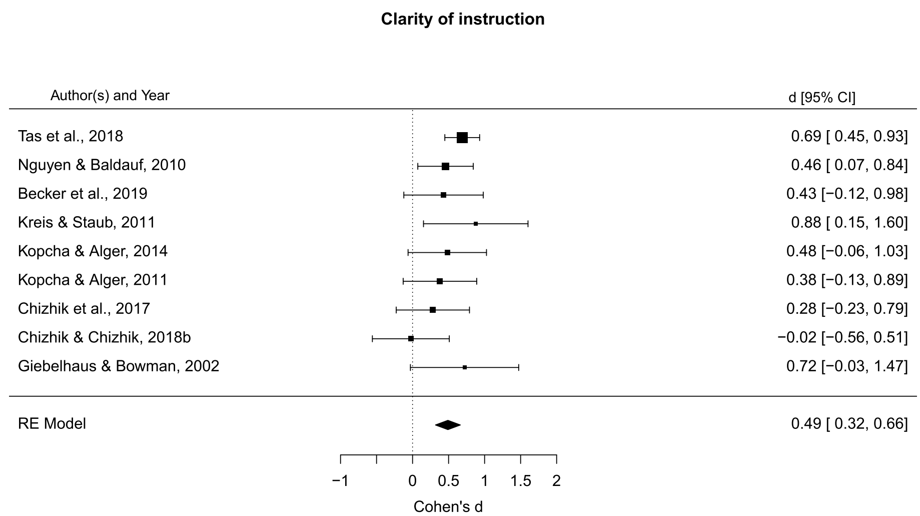

Fig. 3. Overall effect for clarity of instruction ($k = 9$ studies).

Table 2
Moderator effects of learning opportunities for instructional skills of pre-service teachers.

	d	SE	z	p	CI.lb	CI.ub	k	Q
Dialogues during pre-lesson conference				**.742**				0.11
No pre-lesson dialogues *	0.46	.20	2.30	.022	.07	.86	5	
Pre-lesson dialogues *	0.38	.17	2.15	.032	.03	.72	7	
Joint responsibility during dialogues				**.714**				0.13
No joint responsibility †	0.49	.26	1.91	.056	-.01	1.00	3	
Joint responsibility *	0.38	.15	2.50	.013	.08	.69	9	
Cognitive modeling *				**.033**				4.53
No cognitive modeling*	0.28	.13	2.19	.029	.03	.54	9	
Cognitive modeling***	0.90	.26	3.46	<.001	.39	1.41	3	
Personal observation				**.638**				0.22
No personal observation	0.32	.23	1.38	.169	-.14	.78	4	
Personal observation **	0.45	.16	2.90	.004	.15	.76	8	
Opportunities for reflection for PSTs after the lesson				**.60**				0.27
No reflection opportunities †	0.34	.19	1.84	.066	-.02	.71	6	
Reflection opportunities *	0.48	.19	2.58	.010	.12	.85	6	
Receiving structured feedback (after observing a PTS's lesson)				**.74**				0.11
No structured feedback *	0.38	.17	2.15	.032	.03	.72	7	
Structured feedback *	0.46	.20	2.30	.022	.07	.86	5	

Note. $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$, $^{\dagger}p < .10$. The moderator analysis followed two steps. First, an omnibus test indicates if the moderator has a significant influence on the outcome (i.e., PSTs' instructional skills) and if the categorical levels of a moderator differ significantly. Second, the effect size and significance are calculated post hoc for each moderator level.

Table 3
Moderator effects of digital tools and peer coaching for instructional skills of pre-service teachers.

	d	SE	z	p	CI.lb	CI.ub	k	Q
Digital Tools				**.64**				0.21
No digital tools *	0.36	.17	2.12	.034	.03	.70	8	
Digital tools *	0.49	.21	2.33	.020	.08	.90	4	
Peer coaching				**.115**				2.49
No peer coaching ***	0.50	.13	3.76	<.001	.24	.77	10	
Peer coaching	0.04	.26	.16	.871	-.47	.55	2	

Note. $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$, $^{\dagger}p < .10$. The moderator analysis followed two steps. First, an omnibus test indicates if the moderator has a significant influence on the outcome (i.e., PSTs' instructional skills) and if the categorical levels of a moderator differ significantly. Second, the effect size and significance are calculated post hoc for each moderator level.

teachers) struggle not only with their teaching practice (Poulou, 2007), but also with emotional and psychological (e.g., Beck & Kosnik, 2000) as well as motivational aspects of teaching (e.g., self-efficacy, see Holzberger et al., 2013; Woolfolk Hoy & Burke-Spero, 2005). It has repeatedly been shown that mentoring beginning teachers can provide emotional and psychological support (for a review, see Hobson et al., 2009). Although scarce, there is research on the effects of mentoring beginning teachers on self-reported growth in efficacy, job satisfaction, enthusiasm, and reduction of emotional exhaustion (Richter et al., 2013). While our meta-analysis shows that collaboration with trained CTs can improve PSTs' instructional skills, it does not consider CTs' role in providing emotional and psychological support and their impact on fostering emotional and motivational aspects of teaching. Addressing these aspects are relevant because, for example, self-efficacy has been shown to predict evaluated teaching performance (Heneman, Kimball, & Milanowski, 2006; Klassen & Tze, 2014) and self-reported instructional practices (Depaepe & König, 2018; Holzberger & Prestele, 2021). There is also evidence that teachers' reported emotion regulation and psychological need satisfaction can predict motivating teaching style (Moè & Katz, 2020). Unfortunately, research that explores the effects of instructional skills (i.e., lesson planning or clarity of instruction) in the context of mentoring, coaching, or supervision on PSTs' motivational (self-efficacy; e.g., Rupp & Becker, 2021) and emotional aspects of teaching is still scarce. Thus, future studies on PSTs' professional development through mentoring, coaching, or supervision should aim to simultaneously investigate effects on instructional, motivational, and emotional aspects of teaching and their interrelations.

From a methodological point of view, including only quasi-experimental and experimental studies with a control group design is in line with the recommendation by Smith and Ingersoll (2004) to focus more on randomized or quasi-experimental studies with a control group design to support more evidence-based conclusions about effective mentoring, coaching, and supervision. Additionally, these kinds of quantitative studies increase the internal validity of our findings, as changes in outcomes can be clearly traced back to the intervention (Cook et al., 2002; Marsden & Torgerson, 2012). However, readers need to be cautious about generalizing the overall effect we observed on PSTs' instructional skills due to the heterogeneous foci in the teaching practicum across and within countries (Ronfeldt & Reininger, 2012; Ulrich et al., 2020) and due to the diverse challenges CTs face during the practicum in different countries (Clarke & Mena, 2020). Future studies investigating the same teaching practicum foci within countries would be necessary to more fully understand the effects of the practicum.

4.2. Overall effect on PSTs' lesson planning skills and clarity of instruction

As our studies measured two different instructional skills, we also discuss the overall effects on lesson planning and clarity of instruction separately in the following. Our findings showed that PSTs attained gains in lesson planning skills (i.e., a small but not significant overall effect) and the provision of clear instructions (small-to-medium overall effect) when they were supported by trained coaches, mentors, or supervisors during the practicum. The findings are important for multiple reasons. Studies frequently report that unexperienced PSTs show the tendency to use general and naïve knowledge rather than theory-based knowledge to plan their lessons (e.g., Hammerness, Darling-Hammond, & Shulman, 2002). Our overall small but non-significant effect on *lesson planning skills* suggests that CTs or supervisors in the experimental groups in coaching, mentoring, and supervision studies might have been using more sophisticated knowledge and teaching methods that supported PSTs' lesson planning skills in comparison to the control groups. This non-significant effect might be due to the lack of opportunities for PSTs to reflect on the underlying theories behind the teaching practices; PSTs may not have experienced teaching before the practicum, so they may struggle to anticipate concrete practices for future lessons. Lesson planning entails a mental simulation of future teaching actions, but it cannot be concluded that greater lesson planning skills can be directly translated into better teaching practices (Seidel et al., 2013). Therefore, lesson planning is a good initial learning opportunity especially for unexperienced PSTs at the beginning of their practicum (Fuller & Bown, 1975), but it is further essential that PSTs practice their teaching practice during the lesson with assistance and feedback from an experienced CT in real classrooms (Darling-Hammond, 2014). Our results also showed that PSTs attained gains in clarity of instruction (small-to-medium overall effect), consistent with previous research (König et al., 2017). This significant effect size is pivotal because it is not easy for PSTs to provide clear instructions during the practicum even though they know that clear instruction can support students' learning (König et al., 2017). Moreover, our overall effect is based on studies using different instruments that were validated and used in wide-ranging environments to assess teaching-related skills (Danielson, 2013; van de Grift, 2014), and can therefore be considered robust.

4.3. Moderator effects

Learning opportunities. The only significant moderator in our full sample for instructional skills was *cognitive modeling*, which is a pivotal element of cognitive apprenticeship. PSTs who were paired with CTs or supervisors who used cognitive modeling to demonstrate teaching-related practices and explain the reasoning underlying the design of the lesson plan or the selection of teaching practices showed significantly higher instructional skills (large effect) than PSTs who were paired with CTs or supervisors who did not apply cognitive modeling (small effect). This finding is in line with the results of a recent qualitative review revealing that modeling, including making links between theory and practice, is one of the key elements of a quality CT working with PSTs (Ellis et al., 2020). Based on the significant effect of cognitive modeling, we recommend that PST education programs integrate cognitive modeling opportunities for CTs and supervisors when designing their support programs to improve the instructional skills of PSTs in their practicum effectively.

Research by Gibbons and Cobb (2017) suggested different coaching elements as productive for teachers' learning; with our moderator analysis we can only confirm one statistically relevant moderator (i.e., cognitive modeling) for PSTs' instructional skills. Other moderators, such as observing and providing feedback, were not significant in our sample. There could be various explanations for these non-significant moderator effects. The most obvious reason is our relatively small sample size of 12 studies after the exclusion of two outliers. Smaller sample sizes result in low statistical power to find significant differences between subgroups (Hedges & Pigott, 2004). Furthermore, nearly all assisting approaches used a different combination of learning opportunities, which might make it challenging to disentangle individual moderator effects. Thus, if more experimental studies with a clear emphasis on particular learning opportunities are available in the future, scholars will be better able to compare combinations of moderators and single moderators to detect moderator effects of PSTs' instructional skills.

Although the majority of the moderators in our analysis did not indicate significant differences between the subgroups, most single subgroups were consistent with the theoretical assumptions regarding the learning opportunities that are most productive for PSTs' learning (Ellis et al., 2020; Gibbons & Cobb, 2017; McDonald et al., 2013). For example, the subgroup of PSTs who had the opportunity to reflect on their teaching practices showed a significant small-to-medium effect on instructional skills, even though the moderator *reflection* was not significant.

Peer coaching and digital tools. Despite the suggested positive influence of peer coaching (Lu, 2010) and the use of digital tools on instructional skills by previous research (e.g., Kopcha & Alger, 2014; Price & Chen, 2003), we did not find evidence in favor of these new developments in coaching, mentoring, and supervision (Table 3). The moderator *peer coaching* was not significant. Conditions with supplementary peer coaching during practicum resulted in only a very small non-significant effect. In contrast, other approaches without peer coaching had a moderate and significant effect on instructional skills. This is not consistent with findings from an earlier review by Lu (2010) reporting a positive influence of peer coaching on students' learning. Lu's review (2010) included different studies from ours; for example, Lu also included studies focused on students with special needs and a study with an extremely large effect size ($d > 2.0$; Bowman & McCormick, 2000) that we excluded from our meta-analysis. It should be noted that in most peer coaching or peer mentoring studies, the CT also offered minimal guidance during the practicum (i.e., in addition to the peer; Bowman & McCormick, 2000; Lu, 2010). Our results showed that peer coaching might not be more effective for PSTs' instructional skills compared to the support given by CTs. However, it should be noted that the results on peer support were based on only two studies in the final sample (after excluding two outliers with strong positive effects) and, therefore, should be interpreted with caution. More quantitative data are needed to make reliable evidence-based conclusions on the effectiveness of coaching or mentoring by peer students.

As few studies have integrated digital tools into their coaching, mentoring, or supervision approaches and as the moderator *use of digital tools* was not significant, we cannot conclude that using digital tools is more effective than not using these elements. In our sample, we descriptively showed that whereas digital tools within coaching, mentoring, or supervision for PSTs produced a small-to-medium effect on instructional skills. In contrast, not using digital tools revealed a small effect, although the difference between subgroups was not significant. Two studies in our meta-analysis used the same research design but differed in their key digital tools (i.e., using private discussion boards between the CT and PSTs and public discussion boards versus only public discussion boards; Kopcha & Alger, 2011; Kopcha & Alger, 2014). Based on this comparison, it was noticeable that the study that used private discussion boards between CTs and PSTs, which allowed a more intensive exchange on PSTs' individual problems, showed significant effects on lesson planning (Kopcha & Alger, 2014), whereas the earlier study using only public discussion boards did not (Kopcha & Alger, 2011). It seems that the effects of digital coaching, mentoring, and supervision were not related to the use of digital tools *per se*, but the individual support by CTs through digital means that promoted personal reflection on individual teaching situations. Thus, we believe that more experimental studies investigating digital tools that offer individual support for PSTs are necessary to draw conclusions about which elements are most effective for enhancing the instructional skills of PSTs.

4.4. Limitations

The present meta-analysis has two limitations. First, we had a small sample size because we investigated only intervention studies that used quasi-experimental or experimental control-group designs, which resulted in a limited number of eligible studies. Many of the existing quantitative studies focused on gains in PSTs' instructional skills during the field experience were survey-based and did not apply a quasi-experimental or experimental study design. It is challenging to recruit participants for quasi-experimental or experimental study in the context of intense professional development in the field, which may be a further reason for the small number of studies in our meta-analysis. Second, it might be that we did not cover all existing empirical studies, despite our extensive literature search. Nevertheless, as empirical studies especially regarding lesson planning are scarce (König et al., 2020), our meta-analysis provides important preliminary evidence from quasi-experimental and experimental studies exploring coaching, mentoring, and supervision intervention effects on PSTs' instructional skills during the practicum.

5. Conclusion

In our meta-analysis, we found evidence that coaching, mentoring, and supervision in which a CT or supervisor assists PSTs were effective at improving PSTs' instructional skills (i.e., lesson planning skills and clarity of instruction) during the practicum. Our results showed a small and significant overall effect across 12 studies using a quasi-experimental or experimental control-group design. The only significant moderator of coaching, mentoring, and supervision on PSTs' instructional skills was cognitive modeling of lesson planning and teaching practices: PSTs working with CTs or supervisors who used cognitive modeling (i.e., providing reasoning for the selected teaching practices and demonstrating practices) attained a large effect, whereas PSTs working with CTs or supervisors who did not use cognitive modeling attained a small effect. Thus, cognitive modeling of lesson planning and teaching practices should be considered when designing initial teacher training programs to improve PSTs' instructional skills during the practicum. As the existing number of quasi-experimental and experimental studies are limited, more quasi-experimental and experimental studies are needed in the future to make evidence-based conclusions about the effectiveness of different learning opportunities to promote PSTs' instructional skills.

Declaration of competing interest

None.

Acknowledgement

Our thanks go to Oana Costache for her support in coding parts of the data and her very helpful comments on an earlier version of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tate.2021.103484.

References[1]

Allen, J. M. (2011). Stakeholders' perspectives of the nature and role of assessment during practicum. *Teaching and Teacher Education, 27*(4), 742–750. https://doi.org/10.1016/j.tate.2010.12.004

Ambrosetti, A., Knight, B. A., & Dekkers, J. (2014). Maximizing the potential of mentoring: A framework for pre-service teacher education. *Mentoring & Tutoring: Partnership in Learning, 22*(3), 224–239. https://doi.org/10.1080/13611267.2014.926662

Arnold, K.-H., Gröschner, A., & Hascher, T. (2014). *Schulpraktika in der Lehrerbildung: Theoretische grundlagen, konzeptionen, Prozesse und effekte. [Pedagogical field experiences in teacher education: Theoretical foundations, programmes, processes, and effects]* (Waxmann).

Baeten, M., & Simons, M. (2016). Innovative field experiences in teacher education: Student-teachers and mentors as partners in teaching. *International Journal of Teaching and Learning in Higher Education, 28*(1), 38–51.

Barnett, M., Keating, T., Harwood, W., & Saam, J. (2002). Using emerging technologies to help bridge the gap between university theory and classroom practice: Challenges and successes. *School Science & Mathematics, 102*(6), 299–313. https://doi.org/10.1111/j.1949-8594.2002.tb17887.x

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., et al. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*(1), 133–180, 10.3102%2F0002831209345157.

(*) Becker, E. S., Waldis, M., & Staub, F. C. (2019). Advancing student teachers' learning in the teaching practicum through content-focused coaching: A field experiment. *Teaching and Teacher Education, 83*, 12–26. https://doi.org/10.1016/j.tate.2019.03.007.

Beck, C., & Kosnik, C. (2000). Associate teachers in pre-service education: Clarifying and enhancing their role. *Journal of Education for Teaching, 26*(3), 207–224. https://doi.org/10.1080/713676888

Bengtsson, J. (1995). What is reflection? On reflection in the teaching profession and teacher education. *Teachers and Teaching, 1*(1), 23–32. https://doi.org/

[1] Studies marked with an asterix (*) are included in the current meta-analysis.

10.1080/13540600950010103
Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2021). *Introduction to meta-analysis*. John Wiley & Sons.
Borko, H., & Mayfield, V. (1995). The roles of the cooperating teacher and university supervisor in learning to teach. *Teaching and Teacher Education, 11*(5), 501–518. https://doi.org/10.1016/0742-051X(95)00008-8
Bowman, C. L., & McCormick, S. (2000). Comparison of peer coaching versus traditional supervision effects. *The Journal of Educational Research, 93*(4), 256–261. https://doi.org/10.1080/00220670009598714
Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32–42.
Bryant, S. E., & Terborg, J. R. (2008). Impact of peer mentor training on creating and sharing organizational knowledge. *Journal of Managerial Issues, 20*(1), 11–21.
Burns, R., Jacobs, J., & Yendol-Hoppey, D. (2016). Preservice teacher supervision within field experiences in a decade of reform. *Teacher Education & Practice, 29*(1), 46–75. https://doi.org/10.1080/01626620.2016.1226203
Butler, B. M., & Cuenca, A. (2012). Conceptualizing the roles of mentor teachers during student teaching. *Action in Teacher Education, 34*(4), 296–308. https://doi.org/10.1080/01626620.2012.717012
Chizhik, E. W., & Chizhik, A. W. (2018a). Using activity theory to examine how teachers' lesson plans meet students' learning needs. *The Teacher Educator, 53*(1), 67–85. https://doi.org/10.1080/08878730.2017.1296913
(*) Chizhik, E. W., & Chizhik, A. W. (2018b). Value of annotated video-recorded lessons as feedback to teacher-candidates. *Journal of Technology and Teacher Education, 26*(4), 527–552 https://www.learntechlib.org/primary/p/182175/.
(*) Chizhik, E. W., Chizhik, A. W., Close, C., & Gallego, M. (2017). SMILE (Shared Mentoring in Instructional Learning Environments). Effectiveness of a lesson-study approach to student-teaching supervision on a teacher-education performance assessment. *Teacher Education Quarterly, 44*(2), 27–47 https://www.jstor.org/stable/90010517.
Clarke, A., & Mena, J. (2020). An international comparative study of practicum mentors: Learning about ourselves by learning about others. *Teaching and Teacher Education, 90*, 103026. https://doi.org/10.1016/j.tate.2020.103026
Clarke, A., Triggs, V., & Nielsen, W. (2014). Cooperating teacher participation in teacher education: A review of the literature. *Review of Educational Research, 84*(2), 163–202. https://doi.org/10.3102/0034654313499618
Cogan, M. L. (1973). *Clinical supervision*. Boston: Houghton Mifflin.
Collins, A., Brown, J. S., & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American Educator, 15*(3), 6–11.
Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453–494). Erlbaum.
Cook, T. D., Campbell, D. T., & Shadish, W. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. MA: Houghton Mifflin Boston.
Costache, O., Becker, E., Staub, F., & Mainhard, T. (2019). Using uncertainty as a learning opportunity during pre-lesson conferences in the teaching practicum. *Teaching and Teacher Education, 86*, 102890. https://doi.org/10.1016/j.tate.2019.102890
Costa, A. L., & Garmston, R. J. (1994). *Cognitive coaching: A foundation for renaissance schools*. Gordon Publishers, Inc. Disabilities.
Danielson, C. (2013). *The framework for teaching evaluation instrument*. Princeton, NJ: The Danielson Group.
Darling-Hammond, L. (2014). Strengthening clinical preparation: The holy grail of teacher education. *Peabody Journal of Education, 89*(4), 547–561. https://doi.org/10.1080/0161956X.2014.939009
Darling-Hammond, L., Chung, R., & Frelow, F. (2002). Variation in teacher preparation: How well do different pathways prepare teachers to teach? *Journal of Teacher Education, 53*(4), 286–302. https://doi.org/10.1177/0022487102053004002
Dennen, V. P., & Burner, K. J. (2008). The cognitive apprenticeship model in educational practice. In M. D. Merrill, J. J. Van Merrienboer, & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (pp. 425–439). Lawrence Erlbaum Associates.
Depaepe, F., & König, J. (2018). General pedagogical knowledge, self-efficacy and instructional practice: Disentangling their relationship in pre-service teacher education. *Teaching and Teacher Education, 69*, 177–190. https://doi.org/10.1016/j.tate.2017.10.003
Douglas, A. S. (2011). The different learning opportunities afforded student teachers in four secondary school subject departments in an initial teacher education school–university partnership in England. *Journal of Education for Teaching, 37*(1), 93–106. https://doi.org/10.1080/02607476.2011.547035
Ellis, N. J., Alonzo, D., & Nguyen, H. T. M. (2020). Elements of a quality pre-service teacher mentor: A literature review. *Teaching and Teacher Education, 92*, 103072. https://doi.org/10.1016/j.tate.2020.103072
Engestrom, Y. (1999). Activity theory and individual and social transformation. In Y. Engestrom, R. Miettinen, & R. M. Punamaki (Eds.), *Perspectives on Activity Theory* (pp. 19–38). New York: Cambridge University Press.
European Commission. (2013). *Supporting teacher competence Development for better learning outcomes*. Brussels: European Commission Education and Training UE.
Flores, M. A. (2016). Teacher education curriculum. In J. Loughran, & M. L. Hamilton (Eds.), *international Handbook of teacher education* (pp. 187–230). Springer.
Fuentes-Abeledo, E.-J., González-Sanmamed, M., Muñoz-Carril, P.-C., & Veiga-Rio, E.-J. (2020). Teacher training and learning to teach: An analysis of tasks in the practicum. *European Journal of Teacher Education, 43*(3), 333–351. https://doi.org/10.1080/02619768.2020.1748595
Fuller, F. F., & Bown, O. H. (1975). Becoming a teacher. In K. Ryan (Ed.), *Teacher education (74th yearbook of the national society for the study of education) part 2* (pp. 25–52). University of Chicago Press.
Futter, K. (2017). *Lernwirksame Unterrichtsbesprechungen im Praktikum. Nutzung von Lerngelegenheiten durch Lehramtsstudierende und Unterstützungsverhalten der Praxislehrpersonen [Effective lesson conferences during the teaching practicum. The use of learning opportunities by student teachers and cooperating teachers' support*. Klinkhardt.
Gibbons, L. K., & Cobb, P. (2017). Focusing on teacher learning opportunities to identify potentially productive coaching activities. *Journal of Teacher Education, 68*(4), 411–425. https://doi.org/10.1177/0022487117702579
Giebelhaus, C. R., & Bowman, C. L. (2002). Teaching mentors: Is it worth the effort? *The Journal of Educational Research, 95*(4), 246–254.
(*) Goker, S. D. (2006). Impact of peer coaching on self-efficacy and instructional skills in TEFL teacher education. *System, 34*(2), 239–254. https://doi.org/10.1016/j.system.2005.12.002.
Gold, Y. (1996). Beginning teacher support: Attrition, mentoring, and induction. In J. Sikula, T. Buttery, & E. Guy (Eds.), *Handbook of research on teacher education* (pp. 548–594). New York, NY: Macmillan.
Goldhammer, R., Anderson, R. H., & Krajewski, R. J. (1969). *Clinical supervision. Special methods for the supervision of teachers*. New York: Holt, Reinhart and Winston.
Goldhammer, R., Anderson, R. H., & Krajewski, R. J. (1993). *Clinical supervision: Special methods for the supervision of teachers* (3rd ed.). New York: Holt, Rinehart and Winston.
van de Grift, W. J. C. M. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. *Educational Research, 49*(2), 127–152. https://doi.org/10.1080/00131880701369651
van de Grift, W. J. C. M. (2014). Measuring teaching quality in several European countries. *School Effectiveness and School Improvement, 25*(3), 295–311. https://doi.org/10.1080/09243453.2013.794845
Grosser-Clarkson, D., & Neel, M. A. (2019). Contrast, commonality, and a call for clarity: A review of the use of core practices in teacher education. *Journal of Teacher Education, 71*(4), 464–476. https://doi.org/10.1177/0022487119880162
Hammerness, K., Darling-Hammond, L., & Shulman, L. (2002). Toward expert thinking: How curriculum case writing prompts the development of theory-based professional knowledge in student teachers. *Teaching Education, 13*(2), 219–243. https://doi.org/10.1080/1047621022000007594
Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education Review, 30*(3), 466–479.
Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487
Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods, 9*(4), 426–445. https://psycnet.apa.org/doi/10.1037/1082-989X.9.4.426.
Heneman, H. G., Kimball, S., & Milanowski, A. (2006). *The Teacher Sense of Efficacy Scale: Validation evidence and behavioral prediction. WCER working paper No. 2006-7*. Madison, WI: Wisconsin Center for Education Research.
Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*(11), 1539–1558. https://doi.org/10.1002/sim.1186
Hobson, A. J., Ashby, P., Malderez, A., & Tomlinson, P. D. (2009). Mentoring beginning teachers: What we know and what we don't. *Teaching and Teacher Education, 25*(1), 207–216. https://doi.org/10.1016/j.tate.2008.09.001
Hoffman, J. V., Wetzel, M. M., Maloch, B., Greeter, E., Taylor, L., DeJulio, S., et al. (2015). What can we learn from studying the coaching interactions between cooperating teachers and preservice teachers? A literature review. *Teaching and Teacher Education, 52*, 99–112. https://doi.org/10.1016/j.tate.2015.09.004
Holzberger, D., & Prestele, E. (2021). Teacher self-efficacy and self-reported cognitive activation and classroom management: A multilevel perspective on the role of school characteristics. *Learning and Instruction, 101513*. https://doi.org/10.1016/j.learninstruc.2021.101513
Hong, J. Y. (2010). Pre-service and beginning teachers' professional identity and its relation to dropping out of the profession. *Teaching and Teacher Education, 26*(8), 1530–1543. https://doi.org/10.1016/j.tate.2010.06.003
Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Sage Publications.
de Jager, B., Reezigt, G. J., & Creemers, B. P. M. (2002). The effects of teacher training on new instructional behaviour in reading comprehension. *Teaching and Teacher Education, 18*(7), 831–842. https://doi.org/10.1016/S0742-051X(02)00046-X
Joyce, B. R., & Showers, B. (1981). Transfer of training: The contribution of "coaching. *Journal of Education, 163*(2), 163–172. https://doi.org/10.1177/002205748116300208
Jones, J. (1992). *Praxis III teacher assessment criteria research base*. Princeton, NJ: Educational Testing Service.
Kemmis, S., Heikkinen, H. L. T., Fransson, G., Aspfors, J., & Edwards-Groves, C. (2014). Mentoring of new teachers as a contested practice: Supervision, support and collaborative self-development. *Teaching and Teacher Education, 43*, 154–164.
Klassen, R. M., & Tze, V. M. C. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review, 12*, 59–76. https://doi.org/10.1016/j.edurev.2014.06.001

König, J., Bremerich-Vos, A., Buchholtz, C., Fladung, I., & Glutsch, N. (2020). Pre–service teachers' generic and subject-specific lesson-planning skills: On learning adaptive teaching during initial teacher education. *European Journal of Teacher Education, 43*(2), 131–150. https://doi.org/10.1080/02619768.2019.1679115

König, J., Bremerich-Vos, A., Buchholtz, C., Lammerding, S., Strauß, S., Fladung, I., et al. (2017). Modelling and validating the learning opportunities of preservice language teachers: On the key components of the curriculum for teacher education. *European Journal of Teacher Education, 40*(3), 394–412. https://doi.org/10.1080/02619768.2017.1315398

(*) Kopcha, T. J., & Alger, C. (2011). The impact of technology-enhanced student teacher supervision on student teacher knowledge, performance, and self-efficacy during the field experience. *Journal of Educational Computing Research, 45*(1), 49–73. https://doi.org/10.2190/EC.45.1.c.

(*) Kopcha, T. J., & Alger, C. (2014). Student teacher communication and performance during a clinical experience supported by a technology-enhanced cognitive apprenticeship. *Computers & Education, 72*, 48–58. https://doi.org/10.1016/j.compedu.2013.10.011.

Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research, 88*(4), 547–588. https://doi.org/10.3102/0034654318759268

Kreis, A. (2019). Content-Focused Peer Coaching – facilitating student learning in a collaborative way. In T. Janík, I. M. Dalehefte, & S. Zehetmeier (Eds.), *Supporting teachers: Improving instruction. Examples of research-based in-service teacher education*. Waxmann.

(*) Kreis, A., & Staub, F. C. (2011). Fachspezifisches unterrichtscoaching im praktikum [Content-focused coaching in the teaching practicum: A quasi-experimental intervention study]. *Zeitschrift für Erziehungswissenschaft, 14*(1), 61. https://doi.org/10.1007/s11618-011-0170-y.

Kyriakides, L., Christoforou, C., & Charalambous, C. Y. (2013). What matters for student learning outcomes: A meta-analysis of studies exploring factors of effective teaching. *Teaching and Teacher Education, 36*, 143–152. https://doi.org/10.1016/j.tate.2013.07.010

Lawson, T., Çakmak, M., Gündüz, M., & Busher, H. (2015). Research on teaching practicum – a systematic review. *European Journal of Teacher Education, 38*(3), 392–407. https://doi.org/10.1080/02619768.2014.994060

Le Cornu, R., & Ewing, R. (2008). Reconceptualising professional experiences in pre-service teacher education…reconstructing the past to embrace the future. *Teaching and Teacher Education, 24*(7), 1799–1812. https://doi.org/10.1016/j.tate.2008.02.008

Lee, G. C., & Wu, C. C. (2006). Enhancing the teaching experience of pre-service teachers through the use of videos in web-based computer-mediated communication (CMC). *Innovations in Education & Teaching International, 43*(4), 369–380. https://doi.org/10.1080/14703290600973836

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE Publications.

(*) Liu, T.-C. (2005). Web-based cognitive apprenticeship model for improving pre-service teachers' performances and attitudes towards instructional planning: Design and field experiment. *Journal of Educational Technology & Society, 8*(2), 136–149.

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

Lüdecke, D. (2018). *Effect size computation for meta analysis*. https://CRAN.R-project.org/package=esc.

Marcos, J. J. M., Miguel, E. S., & Tillema, H. (2009). Teacher reflection on action: What is said (in research) and what is done (in teaching). *Reflective Practice, 10*(2), 191–204. https://doi.org/10.1080/14623940902786206

Marsden, E., & Torgerson, C. J. (2012). Single group, pre- and post-test research designs: Some methodological concerns. *Oxford Review of Education, 38*(5), 583–616. https://doi.org/10.1080/03054985.2012.731208

Maulana, R., Helms-Lorenz, M., & van de Grift, W. (2015). Development and evaluation of a questionnaire measuring pre-service teachers' teaching behaviour: A rasch modelling approach. *School Effectiveness and School Improvement, 26*(2), 169–194. https://doi.org/10.1080/09243453.2014.939198

McDonald, M., Kazemi, E., & Kavanagh, S. S. (2013). Core practices and pedagogies of teacher education: A call for a common language and collective activity. *Journal of Teacher Education, 64*(5), 378–386. https://doi.org/10.1177/0022487113493807

Mena, J., Hennissen, P., & Loughran, J. (2017). Developing pre-service teachers' professional knowledge of teaching: The influence of mentoring. *Teaching and Teacher Education, 66*, 47–59. https://doi.org/10.1016/j.tate.2017.03.024

Metcalf, K. K. (1992). The effects of a guided training experience on the instructional clarity of preservice teachers. *Teaching and Teacher Education, 8*(3), 275–286. https://doi.org/10.1016/0742-051X(92)90026-Y

Moè, A., & Katz, I. (2020). Emotion regulation and need satisfaction shape a motivating teaching style. *Teachers and Teaching*, 1–18. https://doi.org/10.1080/13540602.2020.1777960

Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods, 11*(2), 364–386. https://doi.org/10.1177/1094428106291059

(*) Nguyen, H. T. M., & Baldauf, R. B., Jr. (2010). Effective peer mentoring for EFL pre-service teachers' instructional practicum practice. *Asian EFL Journal, 12*(3), 40–61.

Orland-Barak, L. (2014). Mediation in mentoring: A synthesis of studies in teaching and teacher education. *Teaching and Teacher Education, 44*, 180–188. https://doi.org/10.1016/j.tate.2014.07.011

Pecheone, R. L., & Chung, R. R. (2006). Evidence in teacher education: The performance assessment for California teachers (PACT). *Journal of Teacher Education, 57*(1), 22–36. https://doi.org/10.1177/0022487105284045

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher, 38*(2), 109–119. https://doi.org/10.3102/0013189X09332374

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment scoring System™: Manual K-3*. Baltimore: Paul H. Brookes Publishing.

Poulou, M. (2007). Student-teachers' concerns about teaching practice. *European Journal of Teacher Education, 30*(1), 91–110. https://doi.org/10.1080/02619760600944993

Price, M. A., & Chen, H.-H. (2003). Promises and challenges: Exploring a collaborative telementoring programme in a preservice teacher education programme. *Mentoring & Tutoring: Partnership in Learning, 11*(1), 105–117. https://doi.org/10.1080/1361126032000054844

R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from www.r-project.org.

Richter, D., Kunter, M., Lüdtke, O., Klusmann, U., Anders, Y., & Baumert, J. (2013). How different mentoring approaches affect beginning teachers' development in the first years of practice. *Teaching and Teacher Education, 36*, 166–177. https://doi.org/10.1016/j.tate.2013.07.012

Rock, M., Gregg, M., Gable, R., Zigmond, N., Blanks, B., Howard, P., et al. (2012). Time after time online: An extended study of virtual coaching during distant clinical practice. *Journal of Technology and Teacher Education, 20*(3), 277–304. https://www.learntechlib.org/primary/p/38623/.

Ronfeldt, M., Brockman, S. L., & Campbell, S. L. (2018). Does cooperating teachers' instructional effectiveness improve preservice teachers' future performance? *Educational Researcher, 47*(7), 405–418. https://doi.org/10.3102/0013189X18782906

Ronfeldt, M., & Reininger, M. (2012). More or better student teaching? *Teaching and Teacher Education, 28*(8), 1091–1106. https://doi.org/10.1016/j.tate.2012.06.003

Rupp, D., & Becker, E. S. (2021). Situational fluctuations in student teachers' self-efficacy and its relation to perceived teaching experiences and cooperating teachers' discourse elements during the teaching practicum. *Teaching and Teacher Education, 99*. https://doi.org/10.1016/j.tate.2020.103252

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Temple Smith.

Scheerens, J. (2015). School effectiveness research. In *International encyclopedia of social and behavioral sciences*. Elsevier.

Seidel, T., Blomberg, G., & Renkl, A. (2013). Instructional strategies for using video in teacher education. *Teaching and Teacher Education, 34*, 56–65. https://doi.org/10.1016/j.tate.2013.03.004

Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research, 77*(4), 454–499. https://doi.org/10.3102/0034654307310317

Slavin, R. E. (2008). What works? Issues in synthesizing educational program evaluations. *Educational Researcher, 37*, 5–14. https://www.jstor.org/stable/30133882.

Smith, E. R. (2005). Learning to talk like a teacher: Participation and negotiation in co-planning discourse. *Communication Education, 54*(1), 52–71. https://doi.org/10.1080/03634520500076778

Smith, T. M., & Ingersoll, R. M. (2004). What are the effects of induction and mentoring on beginning teacher turnover? *American Educational Research Journal, 41*(3), 681–714. https://doi.org/10.3102/00028312041003681

Smith, K., & Lev-Ari, L. (2005). The place of the practicum in pre-service teacher education: The voice of the students. *Asia-Pacific Journal of Teacher Education, 33*(3), 289–302.

(*) Smit, R., Rietz, F., & Kreis, A. (2018). What are the effects of science lesson planning in peers? — analysis of attitudes and knowledge based on an actor–partner interdependence model. *Research in Science Education, 48*(3), 619–636. https://doi.org/10.1007/s11165-016-9581-3.

Staub, F. C. (2004). Transforming educational theory into useable knowledge: A case of co-constructing tools for lesson design and reflection. In B. Ralle, & I. Eilks (Eds.), *Quality of practice-oriented research in science education* (pp. 41–52). Aachen, Germany: Shaker.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap*. New York, NY: Free Press.

(*) Tas, T., Houtveen, T., van de Grift, W., & Willemsen, M. (2018). Learning to teach: Effects of classroom observation, assignment of appropriate lesson preparation templates and stage focused feedback. *Studies In Educational Evaluation, 58*, 8–16. https://doi.org/10.1016/j.stueduc.2018.05.005.

Thurlings, M., Vermeulen, M., Bastiaens, T., & Stijnen, S. (2013). Understanding feedback: A learning theory perspective. *Educational Research Review, 9*, 1–15. https://doi.org/10.1016/j.edurev.2012.11.004

Ulrich, I., Klingebiel, F., Bartels, A., Staab, R., Scherer, S., & Gröschner, A. (2020). Wie wirkt das Praxissemester im Lehramtsstudium auf Studierende? Ein systematischer Review. [How Does the Practical Semester in the Teacher Training Program Affect Students? A Systematic Review]. In I. Ulrich, & A. Gröschner (Eds.), *Praxissemester im Lehramtsstudium in Deutschland: Wirkungen auf Studierende* (pp. 1–66). Springer Fachmedien Wiesbaden.

Valencia, S. W., Martin, S. D., Place, N. A., & Grossman, P. (2009). Complex

interactions in student teaching: Lost opportunities for learning. *Journal of Teacher Education, 60*(3), 304–322. https://doi.org/10.1177/0022487109336543
(*) Veenman, S., Denessen, E., Gerrits, J., & Kenter, J. (2001). Evaluation of a coaching programme for cooperating. *Educational Studies, 27*(3), 317–340. https://doi.org/10.1080/03055690120076592.
Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03
Viechtbauer, W., & Cheung, M. W.-L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods, 1*(2), 112–125. https://doi.org/10.1002/jrsm.11
Wang, J., & Odell, S. J. (2002). Mentored learning to teach according to standard-based reform: A critical review. *Review of Educational Research, 72*(3), 481–546. https://doi.org/10.3102/00346543072003481
Weber, K. E., Gold, B., Prilop, C. N., & Kleinknecht, M. (2018). Promoting pre-service teachers' professional vision of classroom management during practical school training: Effects of a structured online- and video-based self-reflection and feedback intervention. *Teaching and Teacher Education, 76*, 39–49. https://doi.org/10.1016/j.tate.2018.08.008
West, L., & Staub, F. C. (2003). *Content-focused coaching: Transforming mathematics lessons*. Portsmouth, NH: Heinemann.
Woolfolk Hoy, A., & Burke-Spero, R. (2005). Changes in teacher efficacy during the early years of teaching: A Comparison of four measures. *Teaching and Teacher Education, 21*, 343–356.
Zeichner, K. (2002). Beyond traditional structures of student teaching. *Teacher Education Quarterly, 29*(2), 59–64.