Connecting feedback to self-efficacy: Receiving and providing peer feedback in teacher education

Christopher Neil Prilop [a,c,*], Kira Elena Weber [d], Frans J. Prins [b], Marc Kleinknecht [a]

[a] *Leuphana University Lüneburg, Institute of Educational Science, Universitätsallee 1, 21335, Lüneburg, Germany*
[b] *Utrecht University, Department of Education, Faculty of Social and Behavioural Sciences, Heidelberglaan 1, 3584 CS, Utrecht, the Netherlands*
[c] *University of Hamburg, Faculty of Education, Von-Melle-Park 8, 20146 Hamburg, Germany*
[d] *Leibniz Institute for Science and Mathematics Education, Olshausenstr. 62, 24118 Kiel, Germany*

A B S T R A C T

Feedback can play a vital role in fostering teacher self-efficacy. Social comparisons and feedback valence (positive vs. negative feedback) are assumed to have a large impact on self-efficacy. Therefore, how pre-service teachers perceive social comparisons and feedback valence in peer feedback and the extent to which pre-service teachers (bachelor/master students) and teacher trainers incorporate comments that can have an impact on self-efficacy into their peer feedback merit investigation. Two studies were conducted. The first showed that peer feedback consisting of a social comparison and with positive feedback valence resulted in greater willingness to improve and positive affect. The second study revealed that teacher trainers' feedback was more specific, whereas bachelor students' feedback contained more social comparisons than did master students' and teacher trainers'. Future research and practical implications are discussed.

1. Introduction

Education research indicates that receiving feedback from fellow teachers, mentors or supervisors can significantly improve teachers' professional competence (Hammerness et al., 2005; Matsko et al., 2018; Tschannen-Moran & McMaster, 2009). Following Kluger and DeNisi's (1996) meta-analysis of feedback interventions, Hattie and Timperley (2007) ascertained that "feedback is effective to the degree to which it directs information to enhance self-efficacy" (p. 95). Self-efficacy as part of teachers' professional competence has been shown to play a crucial role in student achievement and teacher well-being and commitment (Klassen & Tze, 2014; Zee & Koomen, 2016). Teacher self-efficacy can be considered the belief in one's ability to master challenging teaching situations and reach desired teaching goals (Tschannen-Moran & McMaster, 2009).

Feedback is considered a central source of self-efficacy (Bandura, 1994). Various studies (Cone, 2009; Palmer, 2011; Tschannen-Moran & McMaster, 2009) indicated that feedback can play a meaningful role in increasing teachers' self-efficacy. Feedback can be presented in oral, written or computer-mediated form (e.g., Narciss, 2013; Prilop et al., 2021) and can, for example, contain evaluations or suggestions (e.g.,

Sluijsmans, Brand-Gruwel, Van Merriënboer, & Bastiaens, 2003). However, feedback (i.e., evaluations, suggestions) can also comprise comments such as social comparisons or have different valence that are believed to have a strong impact on self-efficacy (Al-Awidi & Alghazo, 2012; Bandura, 1991). Consequently, effective feedback training courses need to be designed. On the one hand, this requires assessing how feedback comments that can have an impact on self-efficacy are perceived by teachers. On the other hand, which comments are currently present in (pre-service) teachers' peer feedback – that is, assessing the quality of pre-service and expert teachers' peer feedback – need to be determined.

Prior research has highlighted two types of feedback comments as playing a decisive role in the development of self-efficacy: feedback valence and social comparisons. Various studies (e.g., Palmer, 2011; Tschannen-Moran & McMaster, 2009) indicate that feedback valence can influence pre-service teachers' self-efficacy to a large degree. Although evidence emphasises positive feedback valence as a requirement to increase self-efficacy of pre-service teachers (e.g., Tschannen-Moran, Woolfolk Hoy, & Hoy, 1998), balanced feedback valence (positive and negative) is valued as the highest quality type of feedback valence in general without, however, providing sufficient evidence,

especially concerning self-efficacy (e.g., Gielen & DeWever, 2015). In other fields of research, social comparisons have been shown to have substantial effect on peoples' self-efficacy (e.g., Bandura, 1991; Chan & Lam, 2010; Miyake & Matsuda, 2002). To date, feedback research has not systematically analysed how feedback valence or social comparisons influence pre-service teachers' self-efficacy or are present as comments of feedback messages.

The perception of elaborate or concise feedback has been systematically analysed (e.g., Berndt, Strijbos, & Fischer, 2018; Strijbos, Narciss, & Dünnebier, 2010), and feedback quality has been determined in terms of content and/or style criteria such as appropriateness, specificity, quality of suggestions, valence of evaluation and use of the first person (e.g., Prilop, Weber, & Kleinknecht, 2019a; Sluijsmans et al., 2003). However, (pre-service) teachers' self-efficacy has not been taken into account to date. Considering the impact self-efficacy has on (pre-service) teachers and students, feedback research should focus on not only general feedback quality but also the quality of feedback comments that affect self-efficacy (feedback[SE]).

Therefore, we first investigated how pre-service teachers perceive types of written feedback comments that are believed to impact self-efficacy (feedback[SE]: social comparison, feedback valence). Second, we analysed the extent to which pre-service teachers' (bachelor and master students) and teacher trainers' written peer feedback includes feedback[SE] comments. With this approach, the present study contributes to expanding insight into (pre-service) teachers' peer feedback and provides a foundation for future research on fostering the quality of pre-service teachers' peer feedback.

2. Literature review

2.1. Self-efficacy

Teachers' professional knowledge forms the basis for mastery of teaching situations (Kunter et al., 2013). It is based on a variety of sources such as pedagogical knowledge, content knowledge and pedagogical content knowledge but also on affective-motivational components such as self-efficacy beliefs (Blömeke, 2014; Shulman, 1987). In general, self-efficacy can be defined as "people's beliefs about their capabilities to exercise control over their own level of functioning and over events that affect their lives" (Bandura, 1991). It is an integral component of Bandura's social cognitive theory of self-regulation (1991), strongly influencing personal thought, affect, motivation, and action. Though teachers' self-efficacy was first defined as an omnibus trait (Rotter, 1966), it is now conceptualised as a task- and domain-specific multidimensional construct (Zee & Koomen, 2016). Consequently, teachers' self-efficacy can vary between teaching practices and domains (e.g., inclusive practices: Malinen et al., 2013; literacy: Tschannen-Moran & Johnson, 2011; science: Riggs & Enochs, 1990).

2.1.1. Teacher self-efficacy

A growing body of research suggests that self-efficacy has comprehensive implications for teachers and their students (Klassen & Tze, 2014; Zee & Koomen, 2016). Concerning teaching quality, Klassen and Tze (2014) found that (pre-service) teachers' self-efficacy had a strong positive effect on evaluated teaching performance. Self-efficacious pre-service teachers are more open-minded concerning new teaching approaches and focus on meeting their students' needs (Woolfolk Hoy & Burke-Spero, 2005). These findings are supported by a large-scale study by Vieluf, Kunter, and van de Vijver (2013). They observed positive correlations between teacher self-efficacy, job satisfaction and effective classroom teaching practices such as structuring, student orientation and enhanced activities. Regarding student achievement, a comprehensive meta-analysis by Zee and Koomen (2016) showed that, in general, (pre-service) teachers' self-efficacy positively predicted students' overall performance at primary, middle and high school level. Though Zee and Koomen's (2016) meta-analysis revealed only a moderate overall effect, this can still result in important differences in student achievement if it is cumulative over time (Coe, 2002; Klassen & Tze, 2014). In their meta-analysis, Zee and Koomen (2016) were also able to show that self-efficacious teachers generally display higher commitment to the teaching profession (e.g., Chesnut & Burley, 2015). Consequently, if teachers lack self-efficacy, this can have a negative effect on their students' self-efficacy, engagement, motivation and achievement (Zee & Koomen, 2016). Apart from negative effects on student achievement, the worst case is that a lack of self-efficacy can result in teacher burnout (e.g., Schwarzer & Hallum, 2008).

2.1.2. Self-efficacy sources and processes

According to Bandura (1994), self-efficacy can be promoted by four sources (mastery experiences, vicarious experiences, verbal persuasion and physiological arousal) with mastery experience and verbal persuasion believed to be the most important sources. Cognitive, motivational and affective processes are triggered and, at the same time, influenced by self-efficacy (Bandura, 1994). The effect of experiences on people's self-efficacy is based on how sources of self-efficacy are cognitively processed (Morris, Usher, & Chen, 2017). Their interpretation of the information leads them to consider courses of action or set goals. Consequently, their perceived self-efficacy determines what course of action they choose or goal they set. When people have a high sense of self-efficacy, they are more willing to improve and set challenging goals for themselves (Bandura, 1991). People's motivation is determined by how they judge activities. Activities from which individuals derive satisfaction because they perceive themselves to be self-efficacious are pursued with greater willingness and interest (Bandura & Schunk, 1981). Their motivation stems from anticipating possible outcomes. Concerning affective processes, the level of people's self-efficacy determines whether they will experience negative affect when faced with stress or difficult situations (Klassen & Durksen, 2014). For people with a low sense of self-efficacy, experiencing negative affect can undermine performance (Bandura, 1977). In complex tasks, strong negative affect can impair the ability to generate and apply alternative strategies of action (Bandura & Jourden, 1991; Cervone et al., 1991).

The reciprocal relationship between self-efficacy, self-efficacy sources and self-efficacy processes has been described as a cyclical process for the teaching profession (e.g., Burton, Bamberry, & Harris-Boundy, 2005; Salanova, Llorens, & Schaufeli, 2011; Tschannen-Moran & McMaster, 2009). When teachers are provided with the feedback (verbal persuasion) of a colleague, cognitive processes come into play in deciding whether the information provided is adequate and useful. At the same time, affective processes are triggered (physiological arousal), which are also interpreted. Hence, the combination of cognitive and affective processes leads either to more self-efficacy or to less self-efficacy (Klassen & Durksen, 2014). At the same time, teachers perceiving themselves as more efficacious show more willingness to improve and set more challenging goals (motivational and cognitive processes). Achieving predetermined challenging goals (mastery experience) entails positive affect (physiological arousal), which further strengthens teachers' self-efficacy. However, when teachers view themselves as having low self-efficacy, failing in a specific teaching situation could also be detrimental to their self-efficacy. In such instances, coping models (i.e., information by peers also having failed at teaching in the past), contrary to mastery models (i.e., peers reporting only best-practice) can be salient factors in sustaining and increasing self-efficacy (Morris & Usher, 2011). As a result, sources and processes of self-efficacy form "a self-reinforcing cycle of either success or failure" (Tschannen-Moran & McMaster, 2009, p. 229). Consequently, self-efficacy can also be assessed via the cognitive, motivational and affective processes it triggers (e.g., Bandura, 2006; Seo & Ilies, 2009; Woodrow, 2011).

2.2. Feedback

Feedback as a type of social persuasion is considered one of the most important determinants of self-efficacy (Kampkuiper, 2015). At the same time, feedback is perceived as one of the most influential factors promoting achievement in various domains (Hattie & Timperley, 2007; Narciss, 2013). It provides individuals with information about their current performance, to enable them to improve their future performance and reach specific standards (Narciss, 2013). Studies concerning expertise have shown that in the "absence of adequate feedback efficient learning is impossible and improvement only minimal even for highly motivated subjects" (Ericsson, Krampe, & Tesch-Römer, 1993, p. 367). Various studies (e.g., Allen & Wright, 2014; Fisher et al., 2011; Matsumura, Garnier, & Spybrook, 2013; Prilop et al., 2021; Sailors & Price, 2015; Tschannen-Moran & McMaster, 2009; Vogt & Rogalla, 2009) have found significant effects of feedback on (pre-service) teacher's professional knowledge, practice and beliefs and, hence, student achievement. Overall, this indicates that only practising does not suffice to develop expert skills.

2.2.1. Teachers' peer feedback quality and differences in expertise

Teachers need to participate in feedback sessions to foster their cognitive, motivational and affective disposition. Consequently, Hammerness et al. (2005) asserted that teachers actively need to seek feedback to develop teaching expertise. Feedback sessions are increasingly conducted in teacher education (Weber, Gold, Prilop, & Kleinknecht, 2018; Joyce & Showers, 2002; Kraft, Blazar, & Hogan, 2018). (Pre-service) teachers are provided with feedback either after actual classroom practice or after specific skills training. Feedback can be provided by peers who have a similar knowledge-base or by experts with knowledge authority (Lu, 2010).

Teachers' peer feedback competence can be defined as the skill to convey an assessment of a peer's teaching practice to stimulate reflection and, thus, professional development (Hammerness et al., 2005; Sluijsmans et al., 2003; Tripp & Rich, 2012). Peer feedback consists of constructively communicating a criteria-based evaluation of a peer's teaching performance that includes identifying possible strengths and weaknesses (Sluijsmans et al., 2003).

To date, differences of peer feedback quality have been systematically investigated by few studies. Prins, Sluijsmans, and Kirschner (2006) analysed what distinguished experts' (i.e., general practitioner trainers) feedback from novices' (i.e., general practitioners in training) feedback in medical education and found that experts made more use of criteria, provided more situation-specific comments and positive remarks and more frequently used a first-person perspective style. Moreover, a considerable number of novices did not ask any reflective questions (59 %) or did not provide alternative suggestions to proceed (44 %). Prilop et al. (2019a) made similar findings in teacher education. They found that feedback provided by experts was more specific, made use of more questions and used the first-person perspective more frequently than did pre-service teachers' feedback at the bachelor level. In fact, pre-service teachers provided little specific description of teaching situations in their feedback and very rarely used activating questions. Contrary to Prins et al. (2006) findings, experts and novices predominantly provided negative feedback. Concerning specificity and questions, Prilop et al. (2019a) concluded that pre-service teachers lack sufficient professional vision to notice and interpret relevant situations. Studies in professional vision have shown that experts are able to explain and evaluate classroom events more comprehensively than novices are (van den Bogert, van Bruggen, Kostons, & Jochems, 2014; Wolff, Van den Bogert, Jarodzka, & Boshuizen, 2015).

2.2.2. Feedback perceptions

Feedback perception can play a crucial role in feedback effectiveness (Strijbos, Pat-El, & Narciss, 2010) as feedback perception can "mediate between the received peer feedback and subsequent performance"

(Huisman, Saab, Van Driel, & Van Den Broek, 2018, p. 958). Ilgen, Fisher, and Taylor (1979) reviewed the processes through which feedback influences behaviour. They argued that feedback needs to be perceived as useful and be accepted by the recipient and that recipients need to be willing to improve their performance. Feedback content and characteristics of the feedback source can have an impact on how feedback is perceived and the effect it has on recipients' self-efficacy–related cognitive, motivational and affective processes (Bandura, 1991; Strijbos, Narciss et al., 2010; Strijbos, Pat-El et al., 2010; Tschannen-Moran & McMaster, 2009).

Research concerning feedback content has produced mixed results. Straub (1997) investigated student perceptions of teacher comments on a writing task. The study revealed that students "favored detailed commentary with specific and elaborate comments" (p. 91) including advice and explanations. In the same vein, Tsui and Ng (2000) were able to show that only specific and elaborate feedback had an effect on performance. School students working on a text revision incorporated feedback only if it was perceived as helpful due to its suggestions and explanations. Two recent studies analysed the effects of different feedback content. Huisman et al. (2018) investigated the perception of peer feedback in an academic writing task. They were able to show that elaborative peer feedback positively correlated with student perceptions of the adequacy of the peer feedback and their willingness to improve. Berndt et al. (2018) analysed the effect of elaborate specific and concise general feedback content. They established that elaborate feedback leads to more positive affect than concise feedback. On the other hand, there is some research indicating that more elaborate feedback is perceived as less useful and leads to negative affect and less willingness to improve (Shute, 2008). Although the previously mentioned study by Berndt et al. (2018) was based on Strijbos, Narciss et al.' (2010, Strijbos, Pat-El et al., 2010) intervention, the latter established partially adverse results. Though elaborate specific feedback was also perceived as more adequate than concise general feedback, it resulted in negative affect – that is, it can debilitate performance (Bandura & Jourden, 1991; Cervone et al., 1991).

2.3. Fostering self-efficacy through feedback

Self-efficacy can be fostered with feedback. Receiving positive feedback that assures people of their abilities results in higher goals and more commitment to goals (Bandura, 1991; Locke & Latham, 1990). People who are socially persuaded of their mastery increase their efforts when encountering obstacles and are more persistent in succeeding (Bandura, 1977). Furthermore, feedback directed at self-efficacy can enhance intrinsic motivation (Bandura, 1991; Collins, 1982). When feedback is informative, it "enhances performance when there is a clear evidence of progress, but it has little effect when there is considerable ambiguity about the effects of one's courses of action" (Bandura, 1991, p. 251). As a result, focussing on accomplishments instead of failures strengthens one's sense of self-efficacy (Bandura, 1977). Furthermore, providing individuals with social comparisons – information that one's teaching practice was equal, inferior or superior to other teachers' performances – can have beneficial effects on self-efficacy if the comparison does not signify inferiority, that is, it is non-threatening (Bandura, 1991).

A variety of studies (e.g., Al-Awidi & Alghazo, 2012; Chong & Kong, 2012; Cone, 2009; Klassen & Durksen, 2014; O'Neil & Stephenson, 2012; Palmer, 2011; Tschannen-Moran et al., 1998; Tschannen-Moran & McMaster, 2009) showed how feedback can foster or harm (pre-service) teachers' self-efficacy. Tschannen-Moran and McMaster (2009) investigated the effects of four different formats of professional development on the self-efficacy beliefs regarding reading instruction of primary teachers and highlighted the importance of feedback in the form of coaching for the development of self-efficacy. Their results indicated that participants who received feedback after their mastery experiences had the strongest increase of self-efficacy. Teachers who did not receive

feedback showed a decrease in self-efficacy. This led Tschannen-Moran and Tschannen-Moran (2011) to advocate a strength-based coaching approach in contrast to a deficit-based approach to coaching. By focussing on positive practices of teachers, strength-based coaching enables coaches to foster teachers' self-efficacy. Cone (2009), O'Neil and Stephenson (2012) and Klassen and Durksen (2014) investigated the impact different sources had on pre-service teachers' self-efficacy. They found that feedback received from mentors or supervisors influenced pre-service teachers' self-efficacy to a large degree. Similarly, in a study by Al-Awidi and Alghazo (2012), pre-service teachers reported that positive feedback by supervisors or cooperating teachers during a practicum led to an increase of their self-efficacy concerning technology-integration. Furthermore, pre-service teachers found feedback containing social comparisons that indicated superiority to cooperating teachers' skills beneficial to their self-efficacy. The impact feedback had was also emphasized by teachers in a study by Palmer (2011). Teachers rated feedback to be the most effective source of self-efficacy. This was also highlighted by teachers in a study by Chong and Kong (2012) who reported that peer feedback fostered their self-efficacy and increased their willingness to improve.

Fostering high self-efficacy beliefs is especially important for novice teachers. While experienced teachers' self-efficacy is largely fostered by mastery experience, feedback has a more substantial impact on novice teachers' self-efficacy beliefs because they do not experience mastery to the same extent during their first steps in the teaching profession (Moulding, Stewart, & Dunmeyer, 2014; Tschannen-Moran & Woolfolk Hoy, 2007). Bandura (1997) argued that people "hold their efficacy beliefs in a provisional status, testing their newly acquired knowledge and skills before raising their judgments of what they are able to do" (p. 83). Therefore, people require "compelling feedback that forcefully disputes the preexisting disbelief in one's capabilities" (p. 82). As teachers' with high self-efficacy show more commitment to teaching (Chesnut & Burley, 2015), experience less stress and burnout (Aloe, Amo, & Shanahan, 2014; Vesely, Saklofske, & Nordstokke, 2014), but rather job satisfaction (Klassen & Chiu, 2010), high quality peer feedback that fosters teachers' self-efficacy is needed in teacher education and professional development. Consequently, (pre-service) teachers, mentor teachers and supervisors need to be trained in providing high quality feedback (Klassen & Durksen, 2014; Knoblock & Wookfolk Hoy, 2008).

2.4. Feedback quality and self-efficacy

Prior studies measuring feedback quality (Prins et al., 2006; Gielen, Peeters, Dochy, Onghena, & Stuyven, 2010; Gielen & DeWever, 2015) largely have been based on a set of general feedback criteria originally suggested by Sluijsmans, Brand-Gruwel, and Van Merriënboer (2002). Although feedback can have a significant effect on the self-efficacy of feedback recipients, comments that can have an effect on self-efficacy have not been focussed on.

Feedback can be considered the main element of the verbal persuasion source of self-efficacy (Bandura, 1977; Morris et al., 2017). A key objective of verbal persuasion is not to convince people they can successfully complete a task but actually to provide them with the means to achieving success (Bandura, 1977) by providing them with constructive feedback (Gielen et al., 2010). The assessor must be able to explain his or her judgements and highlight specific examples (Prins et al., 2006; Gielen & DeWever, 2015). Furthermore, feedback must contain constructive suggestions (alternatives). Alternatives provide learners with information in addition to evaluative aspects. This can include knowledge about task constraints, concepts, mistakes, how to proceed or teaching strategies (Narciss, 2013). Additionally, feedback messages should contain questions (Gielen et al., 2010) that aim to enhance active engagement (Nicol & Macfarlane-Dick, 2006).

Studies have also evaluated feedback quality by assessing feedback valence – that is, whether feedback messages were positive, negative or

balanced (a mix of positive and negative) (e.g., Prins et al., 2006; Gielen & DeWever, 2015). These studies assumed that feedback messages should contain both positive and negative comments to have the largest impact. Positive or negative feedback can have a direct impact on self-efficacy. While Bandura (1991) argued that negative feedback can lead to increased effort by individuals with high self-efficacy, pre-service teachers in a study by Klassen and Durksen (2014) reported that negative feedback lowered their self-efficacy, while positive feedback increased it. This finding is supported by Tschannen-Moran et al. (1998) who found mostly negative, overly harsh and global rather than specific and empathic, verbal persuasion can lead to a decrease of self-efficacy. Affective reactions following positive or negative feedback determine how people behave (Kluger & DeNisi, 1996). Negative affect can limit the focus on personal inefficacy and potentially unsuccessful actions, whereas positive affect can lead to perceptions of high self-efficacy and setting challenging goals (Bandura, 1982). Though a variety of studies in teacher education (e.g., Al-Awidi & Alghazo, 2012; Klassen & Durksen, 2014; Tschannen-Moran et al., 1998) have indicated that positive feedback increases self-efficacy, it has yet to be established whether balanced (a mix of positive and negative) feedback results in more positive affect and willingness to improve and, hence, has a more positive effect on self-efficacy.

Feedback can also contain social comparisons (Morris & Usher, 2011; Tschannen-Moran et al., 1998). For example, after observing a colleague struggling with student interruptions during teaching, a colleague could indicate that she or he has also experienced similar classroom management problems. Providing an account of one's personal difficulties in performing a task can impact the self-efficacy of others (Bandura, 1991; Morris & Usher, 2011). However, social comparisons have positive impact on self-efficacy only if the information conveyed does not suggest inferiority. Relating this to the prior example, the teacher should provide her or his colleague with instances where she or he had similar problems or found it even harder to cope with certain situations (Al-Awidi & Alghazo, 2012; Morris & Usher, 2011). Consequently, this kind of social comparison avoids the "impression that the task is one of interpersonal competition" (Ilgen & Davis, 2000, p. 561). Comparisons to peers can specifically lead teachers at the onset of their careers to believe that they also have the ability to manage challenging teaching situations (Morris & Usher, 2011; Tschannen-Moran et al., 1998). To our knowledge, social comparisons and their effect on feedback perception have not been investigated in feedback studies (e.g., Prins et al., 2006; Gielen et al., 2010; Gielen & DeWever, 2015) although research in other fields indicates that social comparisons significantly impact self-efficacy (Bandura, 1991; Chan & Lam, 2010; Miyake & Matsuda, 2002).

2.5. Research questions

We conducted two studies to investigate what role peer feedback[SE] comments play in peer feedback perception and provision in teacher education. In the first study, we analysed the effect peer feedback[SE] comments (social comparison, feedback valence) have on pre-service teachers' feedback perception (perceived adequacy, willingness to improve, affect). Feedback perceptions were applied as cognitive, motivational and affective processes triggered by self-efficacy. In the second study, we analysed general peer feedback quality and peer feedback[SE] quality of feedback provided by pre-service teachers (bachelor and master level) and experts (teacher trainers).

Consequently, our research questions are as follows:

(1) How does pre-service teachers' perception of peer feedback[SE] comments (social comparison, feedback valence) differ concerning perceived adequacy, willingness to improve and affect?
(2) To what extent does pre-service (bachelor and master level) teacher and teacher trainer peer feedback differ concerning general peer feedback quality (specificity, alternatives,

questions) and peer feedback^SE quality (social comparison, feedback valence)?

With regard to research question 1, we expect pre-service teachers to perceive peer feedback with a social comparison as having more adequacy and entailing more willingness to improve and greater positive affect than peer feedback without a social comparison comment (Bandura, 1994). Peer feedback that entails non-threatening social comparisons (i.e., displaying similar or inferior competence) makes future improvements more attainable because this kind of social comparison enhances status-equality (Al-Awidi & Alghazo, 2012; Ilgen & Davis, 2000; Morris & Usher, 2011). Equally, we assume that positive peer feedback valence leads to more perceived adequacy, willingness to improve and positive affect among pre-service teachers by causing positive physiological arousal and, thus, promoting pre-service teachers' self-efficacy (Bandura, 1994; Chong & Kong, 2012; Klassen & Durksen, 2014; Woolfolk Hoy & Burke-Spero, 2005).

Regarding research question 2, we presume experts provide higher quality general feedback than pre-service teachers due to having more teaching experience, professional knowledge and professional vision (Prilop et al., 2019a, Prins et al., 2006; van den Bogert et al., 2014; Wolff et al., 2015). With respect to feedback^SE quality, the results of prior studies (Prilop et al., 2019a, Prins et al., 2006) have been mixed with regard to feedback valence. Moreover, studies regarding social comparisons in peer feedback are lacking. Therefore, no specific hypotheses were formulated for feedback^SE quality.

3. Methods and results

3.1. Context of the studies

In Germany, teacher education is divided into two phases (for more information, see Arnold, Gröschner, & Hascher, 2014; Cortina & Thames, 2013): a five-year university phase and a one-and-a-half-year induction phase in a school. Only after completing the induction phase can pre-service teachers become fully qualified teachers. During the university phase, pre-service teachers at our institution must complete bachelor's and master's degrees consisting of two of the following subjects: biology, chemistry, English, religion, art, mathematics, music, politics, general studies, and physical education. At our university, pre-service teachers take part in an observational practicum of 3 weeks in the second bachelor semester, a 4-week practicum including teaching four lessons themselves in the fourth bachelor semester and an 18-week practicum including teaching sixty-four lessons themselves in the second master semester. The bachelor's degrees are identical for all students at our institution. For their master's degrees students need to decide whether they want to become primary (year 1–4) or middle school (year 5–10) teachers. However, three of the four modules (e.g., diagnostics, development of educational institutions, educational achievement) students have to enrol in are the same for master's degrees in primary and middle school education. For the only differing module, students of primary school education have to complete a module on children and childhood studies (Transitions), while students of middle school education have to take a module on the development of adolescents (Psychology for change in secondary schools). After university, pre-service teachers must complete the induction phase. They are supervised by mentor teachers and teacher trainers during this time. Teacher trainers are experienced teachers that applied and were chosen for teacher trainer positions for the induction phase. In Germany, headmasters have to evaluate teachers' performance on a regular basis. This evaluation is taken into account when teachers apply for teacher trainer positions. Teacher trainers receive specific training courses on, for example, coaching and feedback (Dickel, 2009). Teacher trainers have two central tasks. On one hand, they provide trainee teachers with seminars on professional knowledge (Shulman, 1987). On the other hand, teacher trainers have to observe the prospective teachers in their classrooms,

provide them with feedback and finally grade their performances.

3.2. Study 1: pre-service teachers' peer feedback perceptions

3.2.1. Sample

All participating students were enrolled in teaching degrees to become either primary or middle school teachers. The participants participated in courses on school communication, teacher knowledge and inclusive education. In total, 128 students were invited to take part in the study. Participation was anonymous and on a voluntary basis. In all, 104 (76.9 % female, M_{Age} = 27.13, SD_{Age} = 4.92) pre-service teachers at the master level took part in the study.

On average, the pre-service teachers' teaching experience included 1–3 h per week during the previous year on top of their practicums (M =1.85 h, SD = 1.17): *1 – only practicums, 2 – practicums + 1−3 h per week/last year, 3 – 4−7 h per week/last year, 4 – 8−15 h per week/last year, 5 – 16+ h per week/last year*). Additionally, pre-service teachers' self-efficacy was assessed. We measured pre-service teachers' self-efficacy with an established questionnaire developed by Schwarzer and Schmitz (1999). Pre-service teachers rated items such as "*Even if there are disturbances in my class, I am confident I will be able to stay calm*" on a four-point Likert scale ranging from 1 = "strongly disagree" to 4 = "strongly agree". The pre-service teachers displayed high self-efficacy (M = 3.05, SD = 0.33).

3.2.2. Measures: perception of feedback

Pre-service teachers' peer feedback perceptions were assessed with the German version of a questionnaire created by Strijbos, Pat-El et al. (2010). The questionnaire measures peer feedback perceptions by using three scales: *perceived adequacy of feedback*, *willingness to improve* and *affect*. These factors can be seen as cognitive, motivational and affective processes triggered by self-efficacy beliefs. As the scales were developed for higher and further education, we adapted two items for the teaching context. For one item, "my revision" was substituted with "my teaching", and for the other item "performance" was changed to "teaching skills". The first scale assessed how pre-service teachers perceive the *adequacy* of peer feedback, that is, the perceived fairness, usefulness and acceptance of peer feedback. Nine items (e.g., "*I would consider this feedback fair*"; "*I would accept this feedback*") were judged on a five-point Likert scale (1 = "disagree", 5 = "agree"). The second scale evaluated pre-service teachers' *willingness to improve* using three items: whether the peer feedback leads to new goals and motivation (e.g., "*I am willing to improve my teaching skills*"). The third scale measured the *affect* of pre-service teachers after receiving peer feedback by applying six items (e.g., "*I would feel offended/confident if I received this feedback*"). Negatively phrased items were recoded. The internal consistency of the scales was satisfactory (*perceived adequacy of feedback*, Cronbach's α = .81; *willingness to improve*, α = .75; *affect,* α = .82).

3.2.3. Procedure

Pre-service teachers at the master level were presented with a classroom video, followed by a standardised peer feedback message and

Classroom video

Peer feedback

Imagine that you have just given this lesson. You sit down with a colleague and receive the following feedback.

(e.g., Positive feedback type: *I think it's good that you talk to individual children directly when they're not focussing on the task. To notice who is inattentive, I always try to position myself in such a way that I can see the entire classroom.*)

Peer feedback perceptions questionnaire
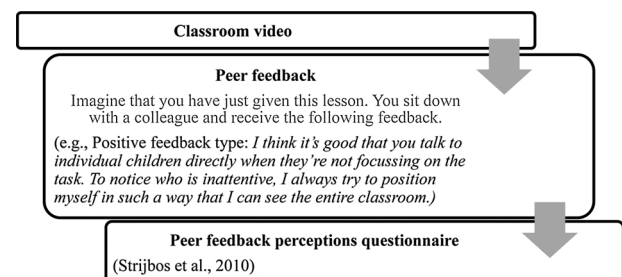(Strijbos et al., 2010)

Fig. 1. Procedure of the study (repeated six times with different peer feedback types).

the peer feedback perceptions questionnaire (see Fig. 1). Pre-service teachers were instructed to imagine they had received the peer feedback message from a colleague after giving the lesson depicted in the video. This procedure was repeated six times in total with different peer feedback types (e.g., positive suggestion with/without social comparison). Participants received a link to the questionnaires at the beginning of the semester and were able to complete the questionnaires online during a course session or at home within the following week. The order of the standardised peer feedback messages was automatically randomised to control for sequence effects.

The one-minute video sequence depicted a classroom management situation. A classroom management situation was chosen as effective classroom management is considered to have significant impact on school students' cognitive, behavioural, and emotional outcomes (Korpershoek, Harms, de Boer, van Kuijk, & Doolaard, 2016. Furthermore, classroom management is a component of (pre-service) teachers' pedagogical knowledge and not content-specific (Shulman, 1987). All pre-service teachers at our university have to complete the same courses covering pedagogical knowledge. Therefore, it can be assumed that pre-service teachers had comparable prior knowledge of classroom management, regardless of the subjects they were enrolled in. The classroom video was recorded in a German primary school in a grade two classroom. The goal of the lesson was that students learned to determine what happens with water when an object is immersed in it. Previously to the scene, the students have noticed that water rises when objects are added. They then formulated assumptions as to what caused the rise in water (e.g., the weight of the object). In the video sequence, the students have come together at different learning stations to assess whether their assumptions are correct. In order to do this, they needed to immerse objects in a plastic cup and jot down their results. The teacher checked the results or asked students to return to their learning stations or to concentrate on the task.

The standardised peer feedback always contained a constructive suggestion ("To notice who is inattentive, I always try to position myself in such a way that I can see the entire classroom") and feedback[SE]. It was either positive ("I think it's good that you talk to individual children directly when they're not focussing on the task"), negative or balanced (a mix of positive and negative). Positive, negative and balanced feedback were accompanied by a social comparison comment or were not. The social comparison comment ("I often experience in my own class that students are inattentive when working at learning stations") signalled that the performance was in line with the colleague's experiences (Bandura, 1991). The pre-service teachers completed the peer feedback perceptions questionnaire for each feedback type.

3.2.4. Methods of analysis

We first analysed differences between pre-service teachers' perception of peer feedback with or without the social comparison (independent variables) comment concerning *perceived adequacy*, *willingness to improve* and *affect* with one-way analyses of variance (dependent variables), followed by Tukey-HSD post hoc tests. We then investigated pre-service teachers' perception of positive, negative or balanced peer feedback (independent variables) regarding *perceived adequacy*, *willingness to improve* and *affect* (dependent variables) with one-way ANOVA. All analyses were computed using SPSS 26 software. The alpha value was set at $p < .05$ for all statistical analyses. No significant departure from a normal distribution of data was detected by the Shapiro-Wilk test. Consequently, the normal distribution of data assumption was met. A priori power analysis using the program GPower showed sufficient statistical power for all analyses (Faul, Erdfelder, Buchner, & Lang, 2009).

3.2.5. Results

Concerning peer feedback types with or without a social comparison (+SC), the ANOVA showed significant differences for *perceived adequacy of feedback*, $F(5,618) = 11.35$, $p < .0001$, *willingness to improve*, F

$(5,618) = 4.26$, $p < .001$, and *affect*, $F(5,618) = 33.36$, $p < .0001$ (see Table 1).

The Tukey-HSD post hoc tests showed that positive feedback + SC resulted in higher positive *affect* than did positive feedback ($p = .015$) but not for *perceived adequacy of feedback* ($p = .589$) or *willingness to improve* ($p = .685$). Balanced feedback + SC did not increase *perceived adequacy of feedback* ($p = .576$), *affect*, ($p = .719$), or *willingness to improve*, ($p = .939$) more than balanced feedback. Negative feedback + SC resulted in higher *perceived adequacy of feedback* ($p < .003$) and higher positive *affect* ($p < .0001$) but not *willingness to improve* ($p = .098$) than did negative feedback. In general, pre-service teachers peer feedback perceptions were higher for feedback with the social comparison comment, concerning *affect* and *perceived adequacy of feedback*, than for feedback without the social comparison comment.

Concerning feedback valence (positive, negative, balanced), the ANOVA showed significant differences for *perceived adequacy of feedback*, $F(2,621) = 18.38$, $p < .0001$, *willingness to improve*, $F(2,621) = 5.70$, $p < .004$, and *affect*, $F(2,621) = 59.293$, $p < .0001$ (see Table 2).

Tukey-HSD post hoc tests showed that positive feedback types (positive and positive + SE) entailed significantly higher perceived adequacy of feedback than negative feedback types (negative and negative + SE; $p < .0001$), but balanced feedback (balanced and balanced + SE) did not ($p = .789$). Balanced feedback also revealed a higher increase than negative feedback ($p < .0001$). No differences were established for *willingness to improve* for positive feedback types in comparison to negative ($p = .07$) or balanced feedback types ($p = .511$). Balanced feedback types increased *willingness to improve* significantly in comparison to negative feedback types ($p = .003$). Positive feedback types increased positive *affect* significantly in comparison to negative feedback types ($p < .0001$) and balanced feedback types ($p < .001$). Balanced feedback types resulted in higher positive *affect* than negative feedback types ($p < .0001$). Overall, positive and balanced feedback types improved perceptions of peer feedback (compared to negative feedback types). Additionally, positive feedback types had a stronger effect on *affect* than did balanced feedback types.

3.3. Study 2: peer feedback quality between career stages

3.3.1. Design and sample

To assess whether teachers' general peer feedback quality and peer feedback[SE] quality differs between career stages, we assessed the quality of feedback of a group of bachelor students ($n = 31$; 80.6 % female; $M_{Age} = 25.45$, $SD_{Age} = 6.01$; $M_{Semester} = 4.57$, $SD_{Semester} = 1.73$), master students ($n = 27$; 74.2 % female; $M_{Age} = 28.15$, $SD_{Age} = 7.00$; $M_{Semester} = 10.04$, $SD_{Semester} = 1.15$) and experts ($n = 20$; 30 % female). The participants produced written feedback for a video-based peer feedback situation. All students were enrolled in teaching programmes to become either primary or middle school teachers. The teacher trainers

Table 1
Means and standard deviations of peer feedback perceptions of peer feedback with and without a social comparison comment.

	Perceived adequacy of feedback		Willingness to improve		Affect	
	M	*SD*	*M*	*SD*	*M*	*SD*
Positive feedback	3.84	0.71	4.09	0.89	3.76	0.68
Balanced feedback	3.89	0.72	4.22	0.79	3.58	0.68
Negative feedback	3.34	0.86	3.83	1.01	2.91	0.76
Positive feedback + SC	4.01	0.82	4.28	0.89	**4.07**	0.65
Balanced feedback + SC	4.07	0.78	4.34	0.86	3.72	0.67
Negative feedback + SC	**3.74**	0.82	4.15	0.93	**3.42**	0.73

Note: Min = 1, Max = 5; significant differences in bold face: with a social comparison > without a social comparison.

Table 2

Means and standard deviations of peer feedback perceptions of peer feedback valence types.

	Perceived adequacy of feedback		Willingness to improve		Affect	
	M	*SD*	*M*	*SD*	*M*	*SD*
Positive feedback types	3.92[a]	0.77	4.18	0.89	3.92[c]	0.68
Balanced feedback types	3.98[b]	0.75	4.28[b]	0.82	3.65[b]	0.67
Negative feedback types	3.54	0.86	3.99	0.98	3.16	0.79

Note: Min = 1, Max = 5; significant difference: [a] = positive feedback valence > negative feedback valence; [b] = balanced feedback valence > negative feedback valence; [c] = positive feedback valence > balanced and negative feedback valence.

(feedback sessions per year: $M = 37.19$, $SD = 33.19$) were experienced teachers responsible for training future teachers in the induction phase. Participants at the bachelor level attended coaching courses connected to a teaching practicum. Participants at the master level took a course on teacher knowledge identical to the one of the participants in study 1, though with a different instructor. The video-based feedback test was conducted in the same way as the questionnaire in study 1. In total, 36 bachelor and 30 master students were invited to take part in the study. Experts were recruited at a professional development course. The video-based test was completed at the beginning of the course in paper–pencil format. The video sequence was presented to the participants with a projector. All the experts participated in the test.

Pre-service teachers at the bachelor level ($M = 1.81$, $SD = 0.70$) and the master level ($M = 2.26$, $SD = 1.53$) did not differ significantly regarding teaching experience ($t(56) = 1.476$, $p = .145$). Experts had a mean of 20.70 years ($SD = 9.48$) of teaching experience. Analysis of pre-service teachers' self-efficacy showed that the bachelor ($M = 3.13$, $SD = 0.31$) and master students ($M = 3.12$, $SD = 0.35$) did not significantly differ ($t(54) = 0.138$, $p = .89$). Pre-service teachers' self-efficacy was measured with the same instrument (Schwarzer & Schmitz, 1999) used in study 1.

3.3.2. Measures: peer feedback quality

Pre-service teachers and experts provided written feedback for a video-based feedback situation. The scenario consisted of information on the teacher, a short video sequence of the teaching situation, a self-reflection and an utterance by the teacher. The content of the scenario was validated concerning the self-efficacy of the teacher and general authenticity. This manipulation check was conducted as the perception of the feedback receiver is believed to have an impact on feedback composition (Narciss, 2013).

The self-reflections and utterances focussed on the aspect of withitness in classroom management. Withitness is a crucial component of effective classroom management. It can be defined as the teacher's ability to continuously monitor what is happening in the classroom and respond appropriately (Kounin, 1970; Wolff, 2015). Self-efficacy determinants were incorporated by adapting items from an established self-efficacy questionnaire (Schwarzer & Schmitz, 1999). These were intended to support an impression of a teacher with low self-efficacy (e. g., "*I don't manage to get the entire class to contribute constructively*"). The classroom video was the same used for the first study.

The impressions of low self-efficacy and authenticity of the peer feedback situation were validated by a different group of experts. They were requested to judge the teachers' self-efficacy in the situation on a six-point Likert scale (1 = "non-existent", 6 = "very high"). In a second step, the experts scored whether the situation was authentic on a four-point Likert scale (1 = "does not apply", 4 = "applies"). The results showed the feedback situation displayed the teacher as lacking self-efficacy ($M = 2.375$, $SD = 0.87$) and that the peer feedback situation was authentic ($M = 3.25$, $SD = 1.04$). Consequently, an acceptable

degree of content validity was provided.

We analysed the written peer feedback concerning the quality of feedback[SE] comments with a newly developed coding scheme (see Table 3) and general peer feedback quality (see Table 4). General peer feedback quality was based on studies by Sluijsmans et al. (2002) and Prins et al. (2006), whereas peer feedback[SE] quality was based on results from our first study. The quality of peer feedback[SE] was assessed as *weakening or no direct effect on self-efficacy* (code 0), *confirming effect on self-efficacy* (code 1) or *strengthening effect on self-efficacy* (code 2). General peer feedback quality was assessed as *absent feedback* (code 0), *average feedback* (code 1) or *good feedback* (code 2).

We analysed the peer feedback, following guidelines for quantifying qualitative data by Chi (1997). Following Prins et al. (2006) coding scheme, which "evaluates the presence of a set of necessary ingredients" (Gielen et al., 2010, p. 307) instead of the frequencies of criteria, comments or words, each feedback was treated as a unit of analysis. The entire message captures the "semantics of the inference at a more appropriate level" (Chi, 1997, p. 10). Before coding the entire sample, 20 randomly chosen feedback comments were coded for calibration. Differences between coders were discussed by the research team and coders to specify the coding manual (Zottmann et al., 2013). Each feedback comment was then independently coded by the coders to establish reliability. As a basic measure of inter-coder reliability, we calculated the percentage of coder agreement, that is, how frequently coders assigned the same code category. The percentage of coder agreement was satisfactory. Due to the possibility of agreement occurring by chance, we additionally calculated Cohen's kappa (κ; Fleiss & Cohen, 1973). The coding yielded satisfactory kappa values (see Tables 3 and 4). Consequently, sufficient inter-coder reliability was established.

3.3.3. Methods of analysis

We applied a one-way multivariate ANOVA (MANOVA), followed by Tukey-HSD post hoc tests of univariate tests, to establish the extent that pre-service teachers' (bachelor and master level) and experts' peer feedback (independent variables) differed concerning general peer feedback quality and peer feedback quality regarding self-efficacy comments (dependent variables). All analyses were computed using SPSS 26 software with an alpha value at $p < .05$ for all statistical analyses. The sample independence assumption was given. Due to the small sample size, a significant departure from a normal distribution of data was detected by the Shapiro–Wilk test ($p < .001$) for all dependent variables. Although normal distribution of data was violated, we still applied parametric analysis tools (MANOVA) as they are robust against this type of violation (O'Brien & Kaiser, 1985; Schmider, Ziegler, Danay, Beyer, & Bühner, 2010). A priori power analysis using the program GPower showed sufficient statistical power for all analyses (Faul et al., 2009).

3.3.4. Results

3.3.4.1. General peer feedback quality and peer feedback[SE] quality.
Concerning research question 2, a MANOVA established large, significant differences between the groups ($p < .001$ (Wilks' lambda), $F (5,10) = 8.42$, $\eta_p^2 = 0.37$). Univariate tests showed that the categories' social comparison ($p = .016$, $\eta_p^2 = 0.10$), specificity ($p < .001$, $\eta_p^2 = 0.50$) and questions ($p = .003$, $\eta_p^2 = 0.14$) differed significantly between groups (see Table 5). Tukey post hoc tests showed that the bachelor students more frequently made social comparisons than did the master students ($p = .047$) and experts ($p = .032$). Master students did not differ significantly from experts ($p = .94$). Concerning the specificity category, experts' feedback displayed significantly more specific situations being elaborated on than did that of bachelor students ($p < .001$) and master students ($p < .001$). Master students' feedback was more specific than bachelor students' ($p < .001$). For the questions category,

Table 3
Content analysis of pre-service teachers' peer feedback[SE] quality: Category, self-efficacy quality, percentage of coder agreement and inter-coder reliability (Cohen's kappa).

Category	Strengthening effect on SE	Confirming effect on SE	Weakening or no direct effect on SE	Percentage of coder agreement	κ
Social comparison	Teaching situation compared to one's own teaching	Teaching situation compared to teaching in general	No comparison of situation	97.4	.64
Valence	Positive evaluation of teaching situation/lesson	Balanced evaluation of teaching situation/lesson	Negative evaluation of teaching situation/lesson	77.9	.67

Table 4
Content analysis of pre-service teachers' general peer feedback quality: Category, feedback quality, percentage of coder agreement and inter-coder reliability (Cohen's kappa).

Category	Good feedback	Average feedback	Absent feedback	Percentage of coder agreement	κ
Specificity	Specific situation elaborated on	Lesson in general elaborated on	No specific situation or lesson in general elaborated on	84.6	.70
Alternatives	Specific alternative presented	General alternative presented	No alternative presented	90.1	.79
Questions	Question explicitly elicits alternative	Question implicitly elicits alternative	No question posed	100	1

Table 5
Content analysis of bachelor and master students' and experts' feedback regarding general peer feedback and peer feedback[SE].

	Bachelor students		Master students		Experts	
	M	*SD*	*M*	*SD*	*M*	*SD*
Social comparison*	0.81	0.83	0.37	0.56	0.30	0.57
Valence	1.84	0.45	2.00	0.00	1.90	0.31
Specificity*	0.97	0.18	1.33	0.55	1.90	0.31
Alternatives	1.65	0.61	1.63	0.74	1.70	0.57
Questions*	0.39	0.67	0.04	0.19	0.65	0.81

Note: Means and standard deviations represent general feedback quality (good feedback, average feedback, absent feedback) or feedback[SE] quality (strengthening effect, confirming effect, weakening or no effect); *: significant differences; Max – 2, Min – 0.

experts' feedback contained significantly more questions than did master students' feedback, ($p = $ .002). Bachelor and master students' quality of questions did not significantly differ ($p < $ .07), and bachelor students' quality of questions did not differ significantly from experts' feedback ($p < $ .28).

4. Discussion

Peer feedback can have a substantial impact on (pre-service) teachers' self-efficacy as verbal persuasion (Bandura, 1994; Hattie & Timperley, 2007). For this reason, the present study first investigated perceptions of different types of peer feedback comments that can have an impact on self-efficacy (feedback[SE]: social comparison, feedback valence) and, second, the extent to which quality criteria concerning general peer feedback quality and peer feedback[SE] quality is found in (pre-service) teachers' peer feedback at different stages of their career (bachelor students, master students, teacher trainers). Concerning the types of peer feedback, ANOVA revealed significant differences. Generally, pre-service teachers' peer feedback perceptions were higher for feedback with a social comparison concerning affect and perceived

adequacy of feedback than for feedback without a self-efficacy comment. Furthermore, ANOVA showed that positive and balanced feedback valence types significantly increased peer feedback perceptions in comparison to negative types of feedback. Additionally, positive feedback valence had a stronger effect on affect than did balanced feedback valence. Regarding the general peer feedback quality and peer feedback[SE] quality comments of bachelor and master students' and experts' feedback, MANOVA indicated that bachelor students' feedback contained more social comparisons than did master students' and expert teachers', experts' feedback was more specific than bachelor and master students', and master students' feedback possessed higher specificity than bachelor students', while expert teachers posed higher quality questions than did master students.

Regarding pre-service teachers' perceptions of peer feedback containing social comparisons, the results support findings of prior studies (Al-Awidi & Alghazo, 2012; Bandura & Barab, 1973; Ilgen & Davis, 2000; Klassen & Durksen, 2014). Peer feedback that displayed social comparisons indicating similar problems resulted in more positive affect than did feedback without a self-efficacy comment. Hence, when teachers receive feedback that shows them that colleagues are struggling with similar problems (coping models), contrary to colleagues only reporting success (mastery models), this leads to positive affect – positive physiological arousal as a source of self-efficacy (Bandura, 1991; Klassen & Durksen, 2014). Providing feedback without a social comparison can result in the feedback being perceived as a demonstration of superiority. Consequently, this would lead to a decrease in self-efficacy by creating negative physiological arousal (Bandura, 1977, 1991). However, voicing that one has also experienced similar problems signals to the feedback recipient that feedback sessions are learning occasions of status-equal participants and not "interpersonal competitions" (Ilgen & Davis, 2000, p. 561). In fact, even negative feedback is more easily dealt with when accompanied by a social comparison, which was illustrated by higher positive affect and perceived adequacy of negative feedback with a social comparison, in contrast to negative feedback without a social comparison.

Concerning pre-service teachers' perceptions of peer feedback valence (positive, negative, balanced), the results confirmed those of prior self-efficacy studies (e.g., Klassen & Durksen, 2014; Tschannen-Moran et al., 1998) that found negative feedback produces the least positive affect and can be detrimental to self-efficacy (Bandura, 1991; Klassen & Durksen, 2014). Our findings indicate that positive feedback has the largest positive impact on affect and, thus, most supports self-efficacy. Again, positive affect leads to positive physiological arousal (Bandura, 1991). This sheds a different light on previous research on peer feedback (e.g., Prins et al., 2006; Gielen & DeWever, 2015) in terms of balanced feedback being assessed as higher quality than purely positive feedback. However, in previous studies, peer feedback quality was measured for a specific task (e.g., feedback reports on consultation skills; feedback on writing a scientific abstract) with a focus on cognitive and meta-cognitive effects and not affective components such as beliefs. Therefore, our results do not necessarily contradict prior studies but instead add to their findings. Balanced feedback might ensure the highest gains when the aim is to foster cognitive, meta-cognitive and affective knowledge components simultaneously.

Our findings partially confirm our hypotheses concerning general

quality components in peer feedback by bachelor and master students and shed some light on how experts and novices take self-efficacy into account. The expert feedback contained higher quality verbal persuasion concerning specificity. This aligns with findings by Prins et al. (2006) and Prilop et al. (2019a). A lack of specificity could be an effect of novices observing classroom events less efficiently than experts (Sherin, Russ, Sherin, & Colestock, 2008; Wolff et al., 2015) and, thus, not perceiving specific situations. This draws attention to the observational aspect of peer feedback. It seems that pre-service teachers' professional vision of classroom management (Gold & Holodynski, 2017) is not sufficiently trained. As classrooms pose highly complex environments, students need support in analysing classroom management situations. An observation sheet could present a viable solution. These provide pre-service teachers with "a particular lens" (Santagata & Angelici, 2010, p. 339) through which to analyse and reflect classroom processes efficiently. Furthermore, expert teachers posed higher quality questions than did master students. This indicates that experts consider feedback a dialogic practice (Hattie & Timperley, 2007; Narciss, 2013). Questions are used to elicit reflection and not to present clear-cut solutions that are open for debate (Tripp & Rich, 2012). Additionally, it appears that participants with especially high self-efficacy do not perceive peer feedback as co-constructive learning opportunities (Prins et al., 2006). Various authors (e.g., Noben, Deinum, Douwes-van Ark, & Hofman, 2021; Wheatley, 2002) propose that a lower degree of self-efficacy might lead to more effective teacher change as teachers then feel a greater psychological need to alter their teaching practice. This might lead to peer feedback being perceived as a more co-constructive learning opportunity. However, considering the impact (pre-service) teachers' self-efficacy can have on (pre-service) teachers and their students (Klassen & Tze, 2014; Zee & Koomen, 2016), we believe (pre-service) teachers need to develop a stronger belief in the value of peer feedback (e.g., Alqassab, Strijbos, & Ufer, 2018). A stronger belief in the value of peer feedback has been associated with more acceptance of feedback (Alqassab et al., 2018). Consequently, providing and receiving peer feedback needs to be more fully incorporated into teacher education curricula, with a focus on feedback as a dialogic practice. Concerning feedback[SE] comments (social comparison, feedback valence), the expert feedback was not of higher quality than the bachelor and master students'. Our results indicate higher quality social comparisons in bachelor students' feedback compared with master students' and experts'. Experts and master students seem to neglect social comparisons. Though social comparisons do not offer specific scaffolds for future practice, as for example alternatives, specificity and questions do, social comparisons can be vital for motivation and persistence (Bandura & Barab, 1973; Morris & Usher, 2011). However, this finding is especially interesting as the bachelor students had not completed any practicums that required them to teach alone. This might indicate that master students, who were evaluated by supervisors during their practicums, and experts, who have to evaluate teachers professionally during the induction phase, do not perceive feedback sessions as learning opportunities but rather as summative assessments, and therefore they do not reveal personal shortcomings (Prins et al., 2006). This finding supports the assertion by prior studies (Knoblauch and Woolfolk Hoy, 2008; Klassen & Durksen, 2014) that mentors or supervisors need to be better trained in providing feedback, especially considering the impact feedback can have on pre-service teachers' self-efficacy. Contrary to extant studies applying tools to assess feedback quality (Prins et al., 2006, Prilop et al., 2019a, 2020), the pre-service teachers predominantly provided peer feedback with positive feedback valence. Previous studies found that positive remarks were more seldom applied by novices in comparison to experts (Prins et al., 2006) or much less than negative remarks (Prilop et al., 2019a, 2020). This indicates that the pre-service teachers in the current study were able to perceive the peer, presented in the video-based feedback situation, as lacking self-efficacy, and they therefore focussed on positive situations. Hence, the design of the feedback situations in previous studies (Prins et al., 2006, Prilop et al., 2019a, 2020) did not offer enough personal information about the peer for novices to make assumptions about the peers' self-efficacy. Prilop, Weber, and Kleinknecht (2019b) were able to show that pre-service teachers are capable of these inferences by analysing the feedback process data of feedback sessions during a teaching practicum. Pre-service teachers provided their peers with more positive feedback than negative feedback.

4.1. Implications for (pre-service) teacher education

Our study established significant differences between expert feedback and pre-service teacher feedback regarding social comparisons. In fact, social comparisons rarely occurred in the master student and expert feedback. This raises concerns regarding the development of peer feedback[SE] competence during career stages as well as the perception of feedback sessions. As a result, there are issues that should be considered in both future research and the implementation of peer feedback in university curricula or professional development courses.

Training courses should be developed to provide pre-service teachers with the opportunity to practice providing feedback[SE] in a less complex environment (Grossman & McDonald, 2008). In designing prospective training courses, distinct prompts (Gielen & DeWever, 2015) should be incorporated to foster peer feedback[SE] (e.g., focussing on successful aspects of the peer's performance or comparing suboptimal aspects of the peer's performance with personal suboptimal experiences). In this context, the use of classroom video can enrich training settings immensely. Video makes it possible to capture the complexity of classroom practice and repeatedly watch segments of a lesson for in-depth analysis (Sherin, 2007) and, due to its proximity to reality, video leads to greater cognitive, emotional and motivational involvement (Seidel, Stürmer, Blomberg, Kobarg, & Schwindt, 2011). Prilop et al. (2020) were able to show that pre-service teachers' general peer feedback quality (specificity, alternatives, questions, use of the first person) can effectively be fostered by training feedback provision based on video sequences of their peers' classroom practice and expert modelling. Training courses should especially focus on social comparisons. The lack of social comparisons in the master students' and experts' feedback seems to indicate that feedback sessions are not perceived as learning opportunities at later career stages. However, increasing knowledge about peer feedback should result in peer feedback sessions being more valued and viewed as status-equal learning opportunities.

Furthermore, future research should assess whether peer feedback has comparable effects on (pre-service) teachers' self-efficacy concerning varying teaching tasks and domains. As teachers' self-efficacy is conceptualised as a multidimensional construct (Zee & Koomen, 2016), feedback might have a stronger impact on (pre-service) teachers' self-efficacy regarding literacy practices (Tschannen-Moran & Johnson, 2011) than on their self-efficacy concerning inclusive practices (Malinen et al., 2013). Moreover, researchers should apply experimental study designs to investigate the effects individual components of general peer feedback quality and peer feedback[SE] quality have on (pre-service) teachers' professional disposition (e.g., knowledge, beliefs, motivation). To provide (pre-service) teachers with authentic experimental conditions, video sequences could be applied as in these studies. This could also yield a more comprehensive picture of which peer feedback comments have the largest impact on self-efficacy. In addition, such studies could also be conducted to identify the most effective social comparison feedback stems. Feedback stems could then be specifically practised in training courses, as outlined above. Moreover, to date, the assessment of self-efficacy relies on self-reports. With new technology such as eye-tracking (e.g., Wolff, 2015) or methods from neuroscience, it may be possible to shed more light on "how beliefs and skills interact in reciprocal fashion" (Schunk, 1998, p. 416). This could also provide more in-depth insight into how peer feedback types specifically impact self-efficacy.

4.2. Limitations

There are some limitations of these studies that should be investigated further in future research. First, we used a sample consisting of master students with high perceptions of self-efficacy to assess the impact of peer feedback[SE] comments (study 1). As one's level of self-efficacy determines whether one experiences negative affect (or not) in challenging situations (Bandura, 1977), this could be the reason pre-service teachers' perceptions of different feedback displayed positive affect. However, even pre-service teachers at the bachelor level with little teaching experience showed similarly high levels of self-efficacy (study 2). Though this might limit our interpretation of the results to pre-service teachers with high self-efficacy, we nevertheless found that different types of peer feedback resulted in differing *perceived adequacy*, *willingness to improve* and *affect*. Second, we used imaginary feedback situations in the studies. Though it might be argued that pre-service teachers would react differently in an actual feedback situation, evidence from other fields of research indicates that imaginary situations have effects comparable to actual situations (e.g., Giacobbe, Stukas, & Farhall, 2013; Papaxanthis, Pozzo, Kasprinski, & Berthoz, 2003; Robinson & Clore, 2001). In study 1, our goal was to investigate how six different feedback types compared to each other such that the comparison effects provided sufficient evidence, and the effect of individual feedback types did not matter. In study 2, the feedback situation was deemed realistic by experts. Overall, our approaches in study 1 and 2, which were based on other studies in peer feedback (Prins et al., 2006; Berndt et al., 2018; Strijbos, Narciss et al., 2010; Strijbos, Pat-El et al., 2010), offered the advantage of allowing the analysis of effects concerning standardised situations. Conducting a study with actual feedback would have been less informative as each teaching situation would have been different, resulting in different feedback. Consequently, we chose this measurement approach as it has proven effective in other studies (Prins et al., 2006, Prilop et al., 2020) and other fields of research (e.g., Heitzmann et al., 2015; Jeschke et al., 2019). Third, we focussed on one component of peer feedback (i.e., self-efficacy). Taking Narciss' (2013) model of conditions and factors influencing feedback processing and effects into account, feedback can be mediated or moderated by conditions and factors relating to the feedback recipient, feedback provider or the feedback situation. Consequently, future studies should analyse how (pre-service) teachers' relationships (mediating or moderating factors of the provider, recipient; e.g., Strijbos, Pat-El, & Narciss, 2021) might affect their processing and provision of feedback[SE]. Such an approach could also provide more comprehensive knowledge on feedback frameworks (e.g., Bell & Cooper, 2013; Kreis & Staub, 2013). In the same vein, we did not assess variables such as professional knowledge or feedback beliefs, which might have had an effect on both the perception of peer feedback (study 1) as well as the composition of written peer feedback (study 2). Regarding the perception of peer feedback, it can be assumed that professional knowledge about classroom management and peer feedback can mediate the *perceived adequacy of feedback*, the *willingness to improve* and positive *affect*. Moreover, beliefs about peer feedback can have an impact on how feedback messages are processed (Strijbos & Müller, 2014). The same argument is valid for the composition of peer feedback. In order to provide peer feedback that has a strengthening effect on self-efficacy, professional knowledge is needed, and participants should hold strong positive beliefs about providing peer feedback (Weaver, 2006). In future research, these variables could be included in order to examine what mediates feedback perceptions and peer feedback quality. Fourth, the participants in our study (study 2) provided written peer feedback. In general, peer feedback is provided orally and face-to-face (Lu, 2010). In our study, participants composed peer feedback in written form, which could have influenced the results due to a lack of verbal (intonation) and non-verbal (facial expression, gesture) information. The differences between verbal feedback and feedback in written form and their impact on feedback perceptions as well as their differences regarding comments that can impact self-efficacy should be investigated more comprehensively.

4.3. Conclusion

Our studies contribute to understanding how peer feedback can best provide messages that enhance self-efficacy (Hattie & Timperley, 2007). The findings highlight the importance of positive peer feedback valence and incorporating social comparisons. Furthermore, the analysis of pre-service teacher and expert feedback showed that there is a need for improvement in both pre-service teachers' and experts' feedback quality, especially concerning feedback comments such as social comparisons. These results can be used in the design of peer feedback training courses (e.g., Alqassab et al., 2018; Sluijsmans et al., 2003), peer feedback frameworks (e.g., Bell & Cooper, 2013; Kreis & Staub, 2013) and in fostering of peer feedback quality (Prilop et al., 2020).

Overall, training and increasing peer feedback quality in pre-service teacher education can lead to (pre-service) teachers engaging more meaningfully with peer feedback. Although feedback sessions are becoming a steady ingredient in (pre-service) teacher education and professional development courses, they are rare in the everyday school environment. Fostering peer feedback quality can cause a change in mindset. Higher quality peer feedback results in greater appreciation of feedback sessions and promotes the development of (pre-service) teachers' competence, including (pre-service) teachers' self-efficacy.

References

Al-Awidi, H. M., & Alghazo, I. M. (2012). The effect of student teaching experience on preservice elementary teachers' self-efficacy beliefs for technology integration in the UAE. *Educational Technology Research and Development, 60,* 923–941.

Allen, J. M., & Wright, S. E. (2014). Integrating theory and practice in the pre-service teacher education practicum. *Teachers and Teaching, 20*(2), 136–151.

Aloe, A. M., Amo, L. C., & Shanahan, M. E. (2014). Classroom management self-efficacy and burnout: A multivariate meta-analysis. *Educational Psychology Review, 26,* 101–126.

Alqassab, M., Strijbos, J.-W., & Ufer, S. (2018). Training peer-feedback skills on geometric construction tasks: Role of domain knowledge and peer-feedback levels. *European Journal of Psychology of Education, 33*(1), 11–30.

Arnold, K.-H., Gröschner, A., & Hascher, T. (2014). *Schulpraktika in der Lehrerbildung: Theoretische Grundlagen, Konzeptionen, Prozesse und Effekte [School practicums during teacher education. Theory, concepts, processes, and effects].* Münster: Waxmann.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84,* 191–215.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122–147.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50,* 248–287.

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp. 71–81). New York: Academic Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: W.H. Freeman.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares, & T. Urdan (Eds.), *Self-Efficacy Beliefs of Adolescents* (pp. 307–337). Greenwich, CT: Information Age Publishing.

Bandura, A., & Barab, P. G. (1973). Processes governing disinhibitory effects through symbolic modeling. *Journal of Abnormal Psychology, 82,* 1–9.

Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology, 41*(3), 586–598.

Bandura, A., & Jourden, F. J. (1991). Self-regulatory mechanisms governing the impact of social comparison on complex decision making. *Journal of Personality and Social Psychology, 60*(6), 941–951.

Bell, M., & Cooper, P. (2013). Peer observation of teaching in university departments: A framework for implementation. *International Journal for Academic Development, 18* (1), 60–73.

Berndt, M., Strijbos, J.-W., & Fischer, F. (2018). Effects of written peer-feedback content and sender's competence on perceptions, performance, and mindful cognitive processing. *European Journal of Psychology of Education, 33*(1), 31–49.

Blömeke, S. (2014). Framing the enterprise: Benefits and challenges of international studies on teacher knowledge and teacher beliefs – Modeling missing links. In S. Blömeke, F.-J. Hsieh, G. Kaiser, & W. H. Schmidt (Eds.), *International perspectives on teacher knowledge, beliefs and opportunities to learn – TEDS-M results* (pp. 3–19). Dordrecht: Springer.

Burton, J. P., Bamberry, N. J., & Harris-Boundy, J. (2005). Developing personal teaching efficacy in new teachers in university settings. *Academy of Management Learning & Education, 4*, 160–173.

Cervone, D., Jiwani, N., & Wood, R. (1991). Goal setting and the differential influence of self-regulatory processes on complex decision-making performance. *Journal of Personality and Social Psychology, 61*, 257–266.

Chan, J. C.-Y., & Lam, S.-F. (2010). Effects of different evaluative feedback on students' self-efficacy in learning. *Instructional Science, 38*(1), 37–58.

Chesnut, S. R., & Burley, H. (2015). Self-efficacy as a predictor of commitment to the teaching profession: A meta-analysis. *Educational Research Review, 15*, 1–16.

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning Sciences, 6*(3), 271–315.

Chong, W. H., & Kong, C. A. (2012). Teacher collaborative learning and teacher self-efficacy: The case of lesson study. *The Journal of Experimental Education, 80*(3), 263–283.

Coe, R. (2002). It's the effect size, stupid: What effect size is and why it is important. In *Paper presented at the annual conference of the British Educational Research Association*. September.

Collins, J. L. (1982). Self-efficacy and ability in achievement behavior. In *Paper presented at the annual meeting of the American Educational Research Association*.

Cone, N. (2009). Community-based service-learning as a source of personal self-efficacy: Preparing preservice elementary teachers to teach science for diversity. *School Science and Mathematics, 109*, 20–30.

Cortina, K. S., & Thames, M. H. (2013). Teacher education in Germany. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers. Mathematics teacher Education* (pp. 49–62). Boston, MA: Springer.

Dickel, M. (2009). "Coaching" – Weiterbildung für Fach- und Hauptseminarleiter: Konzeptionelle Ideen für einen didaktischen Orientierungsrahmen [Coaching – Professional development for content and pedagogical teacher trainers: Conceptional ideas for a didactical frame]. *Seminar – Lehrerbildung und Schule, 3*, 131–147.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363–406.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160.

Fisher, D., Frey, N., & Lapp, D. (2011). Coaching middle-level teachers to think aloud improves comprehension instruction and student reading achievement. *The Teacher Educator, 46*(3), 231–243.

Fleiss, J. L., & Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurements, 33*(3), 613–619.

Giacobbe, M. R., Stukas, A. A., & Farhall, J. (2013). The effects of imagined versus actual contact with a person with a diagnosis of schizophrenia. *Basic and Applied Social Psychology, 35*(3), 265–271.

Gielen, M., & DeWever, B. (2015). Structuring peer assessment: Comparing the impact of the degree of structure on peer feedback content. *Computers in Human Behavior, 52*, 315–325.

Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Stuyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction, 20*, 304–315.

Gold, B., & Holodynski, M. (2017). Using digital video to measure the professional vision of elementary classroom management: Test validation and methodological challenges. *Computers & Education, 107*, 13–30.

Grossman, P., & McDonald, M. (2008). Back to the future: Directions for research in teaching and teacher education. *American Educational Research Journal, 45*(1), 184–205.

Hammerness, K. M., Darling-Hammond, L., Bransford, J., Berliner, D. C., Cochran-Smith, M., McDonald, M., & Zeichner, K. M. (2005). How teachers learn and develop. In L. Darling-Hammond, J. Bransford, P. LePage, K. Hammerness, & H. Duffy (Eds.), *Preparing teachers for a changing world: What teachers should learn and be able to do* (pp. 358–389). San Francisco: Jossey-Bass.

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

Heitzmann, N., Fischer, F., Kühne-Eversmann, L., & Fischer, M. R. (2015). Enhancing diagnostic competence with self-explanation prompts and adaptable feedback. *Medical Education, 49*(10), 993–1003.

Huisman, B., Saab, N., Van Driel, J., & Van Den Broek, P. (2018). Peer feedback on academic writing: Undergraduate students' peer feedback role, peer feedback perceptions and essay performance. *Assessment & Evaluation in Higher Education, 43* (6), 955–968.

Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology, 64*(4), 349–371.

Ilgen, D. R., & Davis, C. A. (2000). Bearing bad news: Reactions to negative performance feedback. Applied Psychology. *An International Review, 49*(3), 550–565.

Jeschke, C., Kuhn, C., Lindmeier, A., Zlatkin-Troitschanskaia, O., Saas, H., & Heinze, A. (2019). Performance assessment to investigate the domain specificity of instructional skills among pre-service and in-service teachers of mathematics and economics. *British Journal of Educational Psychology, 89*(3), 538–550.

Joyce, B. R., & Showers, B. (2002). *Student achievement through staff development* (3rd ed.). Alexandria, VA: ASCD.

Kampkuiper, J. P. (2015). *The effect of positive and negative feedback on self-efficacy, cognitive trust and affective trust. Bachelor Thesis*. University of Twente.

Klassen, R. M., & Chiu, M. M. (2010). Effects on teachers' self-efficacy and job satisfaction: Teacher gender, years of experience, and job stress. *Journal of Educational Psychology, 102*, 741–756.

Klassen, R. M., & Durksen, T. L. (2014). Weekly self-efficacy and work stress during the teaching practicum: a mixed methods study. *Learning and Instruction, 33*, 158–169.

Klassen, R. M., & Tze, V. M. C. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review, 12*, 59–76.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254–284.

Knoblauch, D., & Woolfolk Hoy, A. (2008). Maybe I can teach those kids. The influence of contextual factors on student teachers' efficacy beliefs. *Teaching and Teacher Education, 24*, 166–179.

Kounin, J. S. (1970). *Discipline and group management in classrooms*. New York, NY: Holt, Rinchart & Winston.

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.

Kreis, A., & Staub, F. C. (2013). Kollegiales Unterrichtscoaching [Peer coaching of teachers]. In A. Bartz, M. Dammann, S. G. Huber, T. Klieme, C. Kloft, & M. Schreiner (Eds.), *PraxisWissen SchulLeitung, 33* pp. 1–13). Cologne, Germany: Wolters Kluver.

Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology, 105*(3), 805–820.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Eaglewood Cliffs, NJ: Prentice-Hall.

Lu, H.-L. (2010). Research on peer coaching in preservice teacher education – A review of literature. *Teaching and Teacher Education, 26*, 748–753.

Malinen, O.-P., Savolainen, H., Engelbrecht, P., Nel, J., Xu, M., & Tlale, D. (2013). Exploring the teacher self-efficacy for inclusive practices in three continents. *Teaching and Teacher Education, 33*, 33–44.

Matsko, K. K., Ronfeldt, M., Nolan, H. G., Klugman, J., Reininger, M., & Brockman, S. L. (2018). Cooperating teacher as model and coach: What leads to student teachers' perception of preparedness? *Journal of Teacher Education*, 1–22.

Matsumura, L. C., Garnier, H. E., & Spybrook, J. (2013). Literacy coaching to improve student reading achievement: A multi-level mediation model. *Learning and Instruction, 25*, 35–48.

Miyake, M., & Matsuda, F. (2002). Effects of generalized self-efficacy and negative social comparison feedback on specific self-efficacy and performance. *Psychological Reports, 90*(1), 301–308.

Morris, D. B., & Usher, E. L. (2011). Developing teaching self-efficacy in research institutions: A study of award-winning professors. *Contemporary Educational Psychology, 36*, 232–245.

Morris, D. B., Usher, E. L., & Chen, J. A. (2017). Reconceptualizing the sources of teaching self-efficacy: A critical review of emerging literature. *Educational Psychology Review, 29*, 795–833.

Moulding, L. R., Stewart, P. W., & Dunmeyer, M. L. (2014). Pre-service teachers' sense of efficacy: Relationship to academic ability, student teaching placement characteristics, and mentor support. *Teaching and Teacher Education, 41*, 60–66.

Narciss, S. (2013). Designing and evaluating tutoring feedback strategies for digital learning environments on the basis of the interactive feedback model. *Digital Education Review, 23*, 7–26.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199–218.

Noben, I., Deinum, J. F., Douwes-van Ark, I. M. E., & Hofman, W. H. A. (2021). How is a professional development programme related to the development of university teachers' self-efficacy beliefs and teaching conceptions? *Studies in Educational Evaluation, 68*, Article 100966.

O'Brien, R. G., & Kaiser, M. K. (1985). MANOVA method for analyzing repeated measures designs: An extensive primer. *Psychological Bulletin, 97*(2), 316–333.

O'Neil, S., & Stephenson, J. (2012). Exploring Australian pre-service teachers sense of efficacy, its sources, and some possible influences. *Teaching and Teacher Education, 28*, 535–545.

Palmer, D. H. (2011). Sources of efficacy information in an inservice program for elementary teachers. *Science Education, 95*, 577–600.

Papaxanthis, C., Pozzo, T., Kasprinski, R., & Berthoz, A. (2003). Comparison of actual and imagined execution of whole-body movements after a long exposure to microgravity. *Neuroscience Letters, 339*(1), 41–44.

Prins, F., Sluijsmans, D., & Kirschner, P. A. (2006). Feedback for general practitioners in training: Quality, styles and preferences. *Advances in Health Sciences Education, 11*, 289–303.

Prilop, C. N., Weber, K., & Kleinknecht, M. (2019a). Entwicklung eines video- und textbasierten Instruments zur Messung kollegialer Feedbackkompetenz von Lehrkräften [Development of a video- and text-based instrument to assess teachers' peer feedback competence]. In T. Ehmke, P. Kuhl, & M. Pietsch (Eds.), *Lehrer. Bildung. Gestalten: Beiträge zur empirischen Forschung in der Lehrerbildung*. Weinheim Basel: Beltz Juvent.

Prilop, C. N., Weber, K. E., & Kleinknecht, M. (2019b). How digital reflection and feedback environments contribute to pre-service teachers' beliefs during a teaching practicum. *Studies in Educational Evaluation, 62*, 158–170.

Prilop, C. N., Weber, K. E., & Kleinknecht, M. (2020). Effects of digital video-based feedback environments on pre-service teachers' feedback competence. *Computers in Human Behavior, 102*, 120–131.

Prilop, C. N., Weber, K. E., & Kleinknecht, M. (2021). The role of expert feedback in the development of pre-service teachers' professional vision of classroom management in an online blended-learning environment. *Teaching and Teacher Education, 99*, Article 103276.

Riggs, I. M., & Enochs, L. G. (1990). Toward the development of an elementary teachers' science teaching efficacy belief instrument. *Science Education, 74*, 625–637.

Robinson, M. D., & Clore, G. L. (2001). Simulation, scenarios, and emotional appraisal: Testing the convergence of real and imagined reactions to emotional stimuli. *Personality and Social Psychology Bulletin, 27*(11), 1520–1532.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs, 80*, 1–28.

Sailors, M., & Price, L. (2015). Support for the Improvement of Practices through Intensive Coaching (SIPIC): A model of coaching for improving reading instruction and reading achievement. *Teaching and Teacher Education, 45*, 115–127.

Salanova, M., Llorens, S., & Schaufeli, W. B. (2011). Yes, I can, I feel good & I just do it! On gain cycles and spirals of efficacy beliefs, affect, and engagement. *Applied Psychology: An International Review, 60*, 255–285.

Santagata, R., & Angelici, G. (2010). Studying the impact of the lesson analysis framework on preservice teachers' abilities to reflect on videos of classroom teaching. *Journal of Teacher Education, 61*(4), 339–349.

Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is it really robust? Reinvestigating the robustness of ANOVA against violations of the normal distribution assumption. *Methodology: European Journal of Research Methods for the Behavioural and Social Sciences, 6*(4), 147–151.

Schunk, D. H. (1998). An educational psychologist's perspective on cognitive neuroscience. *Educational Psychology Review, 10*(4), 411–417.

Schwarzer, R., & Hallum, S. (2008). Perceived teacher self-efficacy as a predictor of job stress and burnout: Mediation analyses. Applied Psychology. *An International Review, 57*, 152–171.

Schwarzer, R., & Schmitz, G. S. (1999). Skala zur Lehrerselbstwirksamkeitserwartung (WIRKLEHR). In R. Schwarzer, & M. Jerusalem (Eds.), *Skalen zur Erfassung von Lehrer- und Schülermerkmalen* (pp. 60–61). Berlin: Freie Universität Berlin.

Seidel, T., Stürmer, K., Blomberg, G., Kobarg, M., & Schwindt, K. (2011). Teacher learning from analysis of videotaped classroom situations: Does it make a difference whether teachers observe their own teaching or that of others? *Teaching and Teacher Education, 27*(2), 259–267.

Seo, M., & Ilies, R. (2009). The role of self-efficacy, goal, and affect in dynamic motivational self-regulation. *Organizational Behavior and Human Decision Processes, 109*(2), 120–133.

Sherin, M. (2007). New perspectives on the role of video in teacher education. In J. Brophy (Ed.), *Advances in research on teaching* (pp. 1–27). Bingley, UK: Emerald.

Sherin, M. G., Russ, R. S., Sherin, B. L., & Colestock, A. (2008). Professional vision in action: An exploratory study. *Issues in Teacher Education, 17*(2), 27–46.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1–22.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*, 153–189.

Sluijsmans, D. M. A., Brand-Gruwel, S., & Van Merriënboer, J. J. G. (2002). Peer assessment training in teacher education: Effects on performance and perceptions. *Assessment and Evaluation in Higher Education, 27*(5), 443–454.

Sluijsmans, D. M. A., Brand-Gruwel, S., Van Merriënboer, J. J. G., & Bastiaens, T. J. (2003). The training of peer assessment skills to promote the development of reflection skills in teacher education. *Studies in Educational Evaluation, 29*(1), 23–42.

Straub, R. (1997). Students' reactions to teacher comments: An exploratory study. *Research in the Teaching of English, 31*(1), 91–119.

Strijbos, J. W., & Müller, A. (2014). Personale Faktoren im Feedbackprozess [Personal factors in the feedback process]. In H. Ditton, & A. Müller (Eds.), *Feedback und Rückmeldungen. Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder [Feedback. Theoretical foundations, empirical results, practical fields of application]* (pp. 83–134). Münster: Waxmann.

Strijbos, J. W., Pat-El, R. J., & Narciss, S. (2021). Structural validitiy and invariance of the feedback perceptions questionnaire. *Studies in Educational Evaluation, 68*, Article 100980.

Strijbos, J. W., Narciss, S., & Dünnebier, K. (2010). Peer feedback content and sender's competence level in academic writing revision tasks: are they critical for feedback perceptions and efficiency? *Learning and Instruction, 20*(4), 291–303.

Strijbos, J. W., Pat-El, R. J., Narciss, S., et al. (2010). Validation of a (peer) feedback perceptions questionnaire. In L. Dirckinck-Holmfeld, V. Hodgson, C. Jones, & M. de Laat (Eds.), *Proceedings of the 7th international conference on networked learning* (pp. 378–386). Aalborg, Denmark: Aalborg University [Electronic Proceedings, ISBN 978-1-86220-86225-2].

Tripp, T. R., & Rich, P. J. (2012). Using video to analyse one's own teaching. *British Journal of Educational Technology, 43*(4), 678–704.

Tschannen-Moran, M., & Johnson, D. (2011). Exploring literacy teachers' self-efficacy beliefs: Potential sources at play. *Teaching and Teacher Education, 27*, 751–761.

Tschannen-Moran, M., & McMaster, P. (2009). Sources of self-efficacy: Four professional development formats and their relationship to self-efficacy and implementation of a new teaching strategy. *Elementary School Journal, 110*(2), 228–245.

Tschannen-Moran, B., & Tschannen-Moran, M. (2011). The coach and the evaluator. *Coaching: The New Leadership Skill, 69*(2), 10.

Tschannen-Moran, M., & Woolfolk Hoy, A. (2007). The differential antecedents of self-efficacy beliefs of novice and experienced teachers. *Teaching and Teacher Education, 23*, 944–956.

Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*(2), 202–248.

Tsui, A. B. M., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Learning, 9*, 147–170.

van den Bogert, N., van Bruggen, J., Kostons, D., & Jochems, W. (2014). First steps into understanding teachers' visual perception of classroom events. *Teaching and Teacher Education, 37*, 208–216.

Vesely, A. K., Saklofske, D., & Nordstokke, D. (2014). EI training and pre-service teacher wellbeing. *Personality and Individual Differences, 65*, 81–85.

Vieluf, S., Kunter, M., & van de Vijver, F. J. R. (2013). Teacher self-efficacy in cross-national perspective. *Teaching and Teacher Education, 35*, 92–103.

Vogt, F., & Rogalla, M. (2009). Developing adaptive teaching competency through coaching. *Teaching and Teacher Education, 25*(8), 1051–1060.

Weaver, M. R. (2006). Do students value feedback? Student perceptions of tutors' written responses. *Assessment & Evaluation in Higher Education, 31*(3), 379–394.

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.

Wheatley, K. F. (2002). The potential benefits of teacher efficacy doubts for educational reform. *Teaching and Teacher Education, 18*(1), 5–22.

Wolff, C. E. (2015). *Revisiting 'withiness': Differences in teachers' representations, perceptions and interpretations of classroom management.* Heerlen, Netherlands: Open University of the Netherlands.

Wolff, C. E., Van den Bogert, N., Jarodzka, H., & Boshuizen, H. (2015). Keeping an eye on learning: Differences between expert and novice teachers' representations of classroom management events. *Journal of Teacher Education, 66*(1), 68–85.

Woodrow, L. (2011). College English writing affect: Self-efficacy and anxiety. *System, 39* (4), 510–522.

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* (4), 343–356.

Zee, M., & Koomen, H. M. Y. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being. *Review of Educational Research, 86*(4), 981–1015.

Zottmann, J. M., Stegmann, K., Strijbos, J.-W., Vogel, F., Wecker, C., & Fischer, F. (2013). Computer-supported collaborative learning with digital video cases in teacher education: The impact of teaching experience on knowledge convergence. *Computers in Human Behaviour, 29*(5), 2100–2108.