UNIVERSITÄT AUGSBURG
Universitätsbibliothek

OPEN ACCESS
OPUS AUGSBURG

Enhancing the peer-feedback process through instructional support: a meta-analysis

Julia Hornstein, Melanie V. Keller, Martin Greisel, Markus Dresel, Ingo Kollar

Angaben zur Veröffentlichung / Publication details:

Hornstein, Julia, Melanie V. Keller, Martin Greisel, Markus Dresel, and Ingo Kollar. 2025. "Enhancing the peer-feedback process through instructional support: a meta-analysis." *Educational Psychology Review* 37 (2): 42. https://doi.org/10.1007/s10648-025-10017-3.

Nutzungsbedingungen / Terms of use:

CC BY 4.0

Dieses Dokument wird unter folgenden Bedingungen zur Verfügung gestellt: / This document is made available under these conditions:
CC-BY 4.0: Creative Commons: Namensnennung
Weitere Informationen finden Sie unter: / For more information see:
https://creativecommons.org/licenses/by/4.0/deed.de

Educational Psychology Review (2025) 37:42
https://doi.org/10.1007/s10648-025-10017-3

META-ANALYSIS

Enhancing the Peer-Feedback Process Through Instructional Support: A Meta-Analysis

Julia Hornstein[1] · **Melanie V. Keller**[2] · **Martin Greisel**[1] · **Markus Dresel**[2] · **Ingo Kollar**[1]

Accepted: 31 March 2025
© The Author(s) 2025

Abstract

Peer-feedback can be an effective method to support learning. However, students often require instructional support to provide and process peer-feedback effectively. Previous research used various types of instructional support to improve the quality of peer-feedback processes and outcomes. Yet, a comprehensive overview over their effects is missing. Therefore, this meta-analysis (based on $N = 32$ studies with $N = 3806$ learners) investigates the effects of different kinds of instructional support (feedback provision vs. feedback reception; content-specific vs. generic) on peer-feedback processes (formulating high-quality feedback messages, or effectively reflecting on the feedback received) and outcomes (subject-matter-related knowledge). Overall, peer-feedback with vs. without instructional support had a substantial positive effect ($g = 0.47$). Furthermore, we found a positive effect of feedback provision support on the quality of feedback provision ($g = 0.72$) and the quality of feedback reception ($g = 0.69$) but not on subject-matter-related knowledge. For feedback reception support, we found no effects on peer-feedback processes and outcomes at all. During feedback provision, content-specific support positively influenced the quality of feedback provision ($g = 0.75$) but not subject-matter-related knowledge, while generic support exerts a positive impact on the quality of feedback provision ($g = 0.70$) and subject-matter-related knowledge ($g = 0.55$). During feedback reception, we again found no significant effects of content-related support and generic support at all. The lack of effects for feedback reception support may be related to the limited number of studies on feedback reception in general. Finally, concrete implications and suggestions for future research are provided.

Keywords Instructional support · Peer-feedback · Meta-analysis · Peer-feedback learning processes · Peer-feedback learning outcomes

Extended author information available on the last page of the article

Published online: 21 April 2025

Problem Statement

Peer-feedback has great potential to enhance students' learning processes and outcomes (e.g., Double et al., 2020; Huisman et al., 2019). However, research has shown that students often have difficulties with how to provide high-quality feedback (Carless & Boud, 2018) as well as with how to process and integrate feedback from their peers in their revisions (e.g., Wu & Schunn, 2020a, 2020b). Therefore, students often require additional instructional support to realize the full potential of peer-feedback (e.g., Lui & Andrade, 2022).

Instructional support in peer-feedback may improve different kinds of peer-feedback processes and outcomes: With respect to peer-feedback processes, it may help students formulate high-quality feedback messages or deeply elaborate on the feedback they receive from their peers (Lui & Andrade, 2022; Patchan & Schunn, 2015). With respect to outcomes, instructional support may help students acquire subject-matter-related knowledge on the topic they discussed with their peers (e.g., Dmoshinskaia et al., 2021).

In this article, we propose two ways to differentiate between different kinds of instructional support for peer-feedback: On the one hand, it is possible to distinguish between instructional support according to the particular *phase* of the peer-feedback process in which they are presented (e.g., feedback provision support vs. feedback reception support). However, little is known about the size of effects of instructional support that is presented in a particular phase on the learning processes of that phase, to what extent their effects also affect other phases of the peer-feedback process, and whether they also positively affect learning outcomes. On the other hand, we propose to distinguish between specific *types* of instructional support (e.g., preparatory activities, rubrics, sentence starters, guiding questions, integrated support and content-specific support, generic support). Even though primary studies can be found that show that all these different kinds of instructional support can have beneficial effects on peer-feedback processes and outcomes (e.g., Altstaedter, 2016; Bürgermeister et al., 2021; Gyamfi et al., 2024; Jurkowski, 2018; Zhao et al., 2023), it is still unclear whether their effects are comparable in size or whether some types of instructional support are more effective than others.

Therefore, our aim is to conduct and present the results of a meta-analytic comparison of the effects of different kinds of instructional support on different peer-feedback processes and outcomes. The results of this article provide a solid basis for practitioners to decide on how to effectively design peer-feedback scenarios for their classrooms, lecture halls, or online learning environments. Also, the results of this meta-analysis may inform future theory-building on the mechanisms by which different kinds of instructional support may affect student learning in peer-feedback scenarios.

What Is Peer-Feedback?

During peer-feedback, learners solve a given task and subsequently provide feedback to each other's task solutions, instead of receiving feedback from a teacher or a computer program. We define peer-feedback based on Huisman et al. (2019) as

"all task-related information that a learner communicates to a peer of similar status which can be used to improve his or her [peer-feedback processes and outcomes]." (p. 865). In line with Huisman et al. (2019), such feedback can come in various types of information. Also, just as Huisman et al. (2019), we use the terms peer-feedback and peer assessment as synonyms during this article. According to Bauer et al. (2023), the peer-feedback process consists of five phases: In the (1) *design* phase, the teacher decides which assignment can be used for further processing, how many learners should provide feedback to each other, from how many peers the learners should receive feedback, and whether and how the learners are supported in the subsequent phases of the peer-feedback process. In the (2) *task performance* phase, the learners produce a first task solution on which their peer(s) can give feedback, such as a written analysis of a specific problem (Kollar & Fischer, 2010). In the (3) *feedback provision* phase, initial task solutions are shared with the selected feedback providers, and the learners provide feedback on each other's task solutions. This is followed by the (4) *feedback reception* phase, during which learners process the feedback they received from their peers, reflect on how to use the feedback to improve their task solutions, and revise their initial task solutions based on the feedback. Lastly, the learners evaluate the learning outcomes and the feedback process during the (5) *evaluation* phase (Bauer et al., 2023).

Previous Meta-Analytical Evidence on the Effectiveness of Peer-Feedback

Several meta-analyses have compared the effects of peer-feedback to other types of feedback or to no feedback at all (e.g., Double et al., 2020; Huisman et al., 2019; Li et al., 2020). Taken together, the results of these meta-analyses demonstrated that peer-feedback yields better outcomes than no peer-feedback or, in some cases, other forms of feedback, including teacher feedback and self-assessment. Furthermore, the meta-analyses (Double et al., 2020; Huisman et al., 2019; Li et al., 2020) showed that different moderator variables can improve the effects of peer-feedback on processes and outcomes. Particularly important in the context of this article are the significant positive effects of rater training for the feedback provider, explicit rating criteria, and rubrics, which suggest that instructional support can play an important role in the peer-feedback process (Double et al., 2020; Li et al., 2020). Thus, the results of the moderator analyses indicated that peer-feedback compared to no peer-feedback can be effective in different contexts, and that instructional support may further boost its effects, but that the effects of different kinds of instructional support could vary across studies.

This is supported by the meta-analysis by Zheng et al. (2020), who compared the effects of technology-facilitated peer-feedback and supplemental instructional support methods in technology-facilitated peer-feedback on learning achievement as part of subject-matter-related knowledge. Their results revealed that technology-facilitated peer-feedback had moderate positive effects on learning achievement

compared to non-technology-facilitated peer-feedback. In addition, the use of supplemental instructional support compared to no instructional support led to medium positive effects on learning achievement. Among the types of instructional support Zheng et al. (2020) investigated, "peer-feedback rules" and "extra tools" were found to have significant and medium-sized positive effects on achievement. Further, Zheng et al. (2020) found in moderator analyses that receiving training during peer-feedback can improve learning.

Despite the important contributions of these previous meta-analyses, however, important questions are still unanswered: First, with respect to the dependent variables that were included, most meta-analytical evidence so far mainly focused on learning outcomes, that is on constructs that are measured in subsequent post-tests, that means *after* the peer-feedback process is completed (e.g., students' subject-matter-related knowledge or learning achievement). However, scores in subsequent post-tests are a rather distal measure for the effects of different kinds of instructional support within peer-feedback. It would therefore be interesting to also look at the effects of instructional support on proximal outcomes that emerge *during* the peer-feedback process (such as the quality by which students provide feedback to their peers, or the extent to which they use the feedback they receive in their revisions). For ease of writing, we will use the term "processes" for these variables that are generated during the peer-feedback process in the remainder of this article, even though they are technically proximal outcomes of processes that occur during feedback provision and feedback reception. Second, with the exception of Zheng et al. (2020), previous meta-analyses have not sufficiently investigated whether different kinds of instructional support might have differential effects on different peer-feedback processes and outcomes. For example, Zheng et al. (2020) only addressed instructional support related to *technology-facilitated* peer-feedback and compared only two very specific kinds of instructional support (i.e., peer-feedback rules vs. extra tools). This leaves it open whether there are further—potentially even more effective—types of instructional support to foster peer-feedback. And third, it remains unclear what specific phase(s) of peer-feedback should best be supported, that is, whether it is more effective to support learners during feedback provision or during feedback reception. We elaborate on this latter issue in the following section.

Feedback Provision and Feedback Reception as Crucial Components of the Peer-Feedback Process

As just mentioned, it is still an open question during which phase of the peer-feedback process learners should ideally be supported. In the context of this article, we focus on (a) the feedback provision phase and (b) the feedback reception phase, as these two phases are at the core of the peer-feedback process and bear tremendous potential to support learning. This section focuses on the mechanisms that take place in these two phases that may influence the effectiveness of peer-feedback and on how instructional support might be used to support students' during these phases.

In general, students engage in both overt and covert processes when providing feedback, that is, visible activities (e.g., formulating a feedback message) and

invisible, cognitive processes (e.g., explaining, evaluating). While *providing* feedback, students engage with and critically reflect on their peers' initial task solution, which may help them take an alternative perspective on how to solve the task at hand. By elaborating on their peer's task solution, and by comparing it with their own (Lundstrom & Baker, 2009; van Popta et al., 2017), they may thus engage in deep cognitive processes that may help them to advance their own understanding (Double et al., 2020; Ion et al., 2018). In addition, when justifying their feedback, students evaluate their peer's initial task solutions and may provide subject-matter-related explanations, which all represent important learning activities in their own right (e.g., Bisra et al., 2018; Lachner et al., 2022). Further specific cognitive processes that occur in the feedback provision phase are comparing and challenging ideas, as well as reflecting upon and generating suggestions for change (e.g., Li & Grion, 2019; van Popta et al., 2017). To do this, students need to plan, monitor, regulate, and reflect their own thinking in order to be able to provide high quality feedback messages (Çevik, 2015; van Popta et al., 2017). In addition, when dealing with the initial task solution of their peers, students also need to relate the aspects mentioned in their peers' initial task solutions to their own level of knowledge and explain their point of view (e.g., Li & Grion, 2019; van Popta et al., 2017). This is only possible if students are prepared to take on other perspectives. Each of these cognitive processes lead to a higher level of critical thinking and reflection (Nicol et al., 2014; van Popta et al., 2017). All these cognitive processes and the associated benefits for the feedback provider take place simultaneously with the formulation of high quality feedback messages.

Unfortunately though, prior research shows that the quality of the feedback messages that peers produce often is rather low (Alemdag & Yildirim, 2022). A low feedback quality is problematic because the feedback receiver should benefit much less likely from low quality feedback than from high quality feedback (e.g., Wu & Schunn, 2020a). Therefore, peer-feedback providers may require instructional support to increase the quality of their feedback messages. In this article, we use the term "instructional support" in a liberal sense, denoting all kinds of support that help learners engage in a practice that would otherwise be out of reach (cf. Davis & Miyake, 2004; Jackson et al., 1994; Simons & Klein, 2007). The literature suggests that students can be supported in this cognitively demanding activity using different kinds of instructional support (e.g., Bürgermeister et al., 2021; Latifi et al., 2021a; Valero et al., 2019). When considering related literature, at least two ways to distinguish between different kinds of instructional support seem to be reasonable: First, different kinds of instructional support come in different presentation modes. For example, they can be designed as preparatory activities (e.g., worked examples, discussions about feedback criteria, training), rubrics (e.g., rating schemes), sentence starters (e.g., "To improve your analysis, you could…"), guiding questions (e.g., "Which comments are correct and which comments are incorrect?"), or integrated support (as a combination of different kinds of instructional support). Second, instructional support can also be categorized with respect to the level of specificity of the dependent variables they focus on (e.g., Belland et al., 2013; Ertl et al., 2007). *Content-specific support* provides learners with explicit support on, for example, whether a certain domain-specific term has been used in a correct way in

a peer's initial task solution. The focus of content-specific support is therefore more on the application of specific knowledge in a selected subject area (e.g., Hannafin et al., 1999; Vogel et al., 2017). *Generic support*, in contrast, assists students during peer-feedback independently from the content, for example, by prompting them to make sure that their feedback reminds their peers of the actual task (i.e., feed up; e.g., "The aim is to…"), describes how they did in their initial task solution (i.e., feed back; e.g., "What you have already done well is…"), or to provide suggestions for improvement (i.e., feed forward; e.g., "For the future, hold on to…"; Bürgermeister et al., 2021; Hattie & Timperley, 2007). That way, the complexity of the task of providing feedback is reduced through breaking it down into smaller aspects to mitigate the danger of potential cognitive overload (Hein & Prediger, 2017; Reiser, 2004; Sweller, 1999). Note that these two categorizations of instructional support (i.e., different presentation modes and level of specificity of the dependent variable) are orthogonal to each other. In other words, instructional support that comes as rubrics, questions, sentence starters, etc., will in some cases be content-specific and in other cases generic. Since little is known about possible differential effects of different kinds of instructional support, we therefore apply both categorizations in this article.

Although these types of instructional support can increase the quality of feedback messages, they do not guarantee that students who receive feedback will actually use it and do so in an appropriate way (e.g., Carless & Boud, 2018). For this reason, also the *feedback reception* phase plays a crucial role in the peer-feedback process. One potential advantage of *receiving* feedback from a peer is that students often find peer-feedback more understandable and helpful than teacher feedback because it is written in simpler language, as it is formulated by a person with a similar level of knowledge (Falchikov, 2013; Li et al., 2020). Again, both overt (e.g., using the feedback to revise) and covert cognitive processes (e.g., critical thinking; mindful cognitive processing) take place. However, learners need to actively engage with the feedback they receive in order to manage emotions, interpret the feedback by meaning making und judging (i.e., is the feedback received accurate and relevant?), and make decisions on how to revise their own initial task solution based on it (e.g., Carless & Boud, 2018; Lui & Andrade, 2022). In this process, they compare the information from the feedback with their own task solution, which may help them expand their existing knowledge or even develop new knowledge (Nicol & McCallum, 2022). Many important cognitive processes also take place at this phase: Students are encouraged to mindfully process the feedback and reflect critically by questioning the feedback they receive from their peers, applying the underlying task criteria, comparing the received feedback with their own initial task solution, and creating a learning transfer (e.g., Bolzer et al., 2015; Nicol et al., 2014; Winstone et al., 2017).

Yet, also during this phase, research has shown that students often have difficulties, for example, differentiating between more and less important passages in the feedback (e.g., Jurkowski, 2018). Thus, learners often benefit from being supported by specific types of instructional support. The goal of instructional support is to improve students' ability to deal with the feedback they receive (e.g., that

they are able to assess whether the feedback they received is relevant and accurate; Jurkowski, 2018) so that they can make effective use of the benefits mentioned above. Here again, it is possible to distinguish between (1) different presentation modes of instructional support (e.g., preparatory activities, rubrics, sentence starters, guiding questions, and integrated support; Wichmann et al., 2018) and (2) the level of specificity of the dependent variables the instructional support focuses on (content-specific vs. generic support; Belland et al., 2013; Ertl et al., 2007). However, an integrative overview over the possibly differential effects of these different kinds of instructional support appears to be missing as of now. This is where our meta-analysis steps in.

Research Questions

Previous research (e.g., Dmoshinskaia et al., 2021; Latifi et al., 2021b) has used a variety of different types of instructional support to promote these peer-feedback processes and outcomes. Even though prior meta-analyses (e.g., Double et al., 2020; Huisman et al., 2019; Li et al., 2020) proved that peer-feedback as an instructional method is effective to support learning per se, the large variability of effect sizes that are reported in these meta-analyses warrants a more detailed account to explore what kinds of instructional support work and for what learning processes and outcomes they are effective. Therefore, this meta-analysis asks the following research questions (for a graphical representation, see Fig. 1):

1. What is the overall effect of instructional support during peer-feedback compared to peer-feedback without instructional support?
2. What are the effects of instructional support (i.e., without differentiating between different kinds of instructional support) during peer-feedback on (a) the quality of feedback provision, (b) the quality of feedback reception, and (c) subject-matter-related knowledge?
3. What are the overall effects of (a) feedback provision support and feedback reception support, (b) specific kinds of instructional support *during peer-feedback provision*, and (c) specific kinds of instructional support *during peer-feedback reception* (i.e., without differentiating between different kinds of dependent variables)?
4. What are the effects of (a) feedback provision support and feedback reception support, (b) specific kinds of instructional support *during peer-feedback provision*, and (c) specific kinds of instructional support *during peer-feedback reception* on (a) the quality of feedback provision, (b) the quality of feedback reception, and (c) subject-matter-related knowledge?

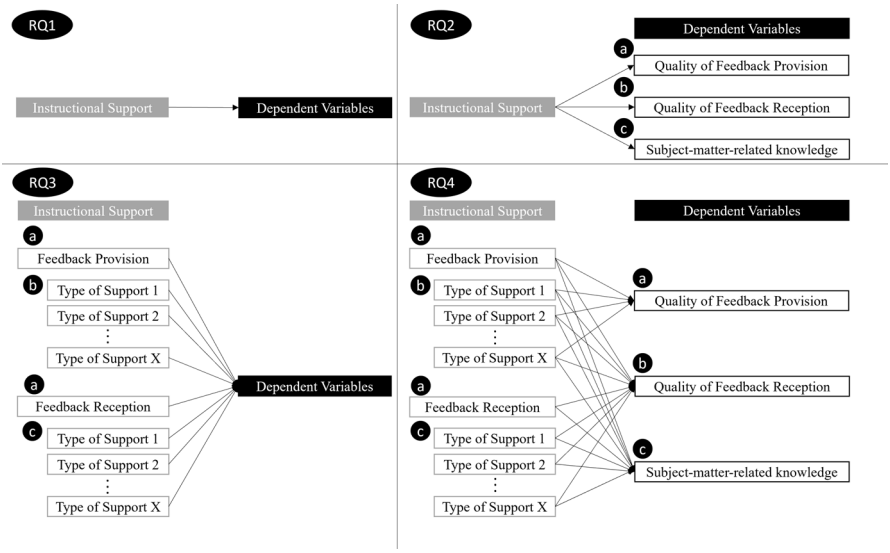

Fig. 1 Graphical representation of research questions

Method

Literature Search

We conducted a systematic literature search in December 2024 using the search term ("peer-feedback" OR "peer assess*" OR "digital feedback" OR "feedback tool" AND ("experimental" OR "quasi experimental" OR "intervention" OR "treatment") in various databases (Web of Science, Psyndex, PubMed, ERIC, PsycArticles, PubPsych, Google Scholar, ProQuest). We imported all hits into the literature management program "Citavi" to sort the records. Google Scholar and ProQuest provided too many hits to import all the entries, so we scanned the first ten results pages for relevance to see if there were additional hits. This database search identified 3065 records. After this, we carried out cross-referencing and a search in ResearchGate, which yielded an additional 128 hits. This allowed us to identify a total of 3193 records (see Fig. 2 for an overview).

The first step was to remove all duplicates, which resulted in 706 records being excluded. The records were then further screened by title and abstract, using this information to determine whether the records were thematically relevant to our research questions. In this step, further 2099 records were excluded. For the remaining 387 records, we scanned the full texts of each article. For records for which we had no access to the full text, we contacted the authors. Since not all authors responded, we had to exclude eight records. In the end, we scanned 379 full texts, of which 32 articles met all inclusion criteria (see below). The first author and the second author performed the screening with a very good interrater reliability (Gwet's $AC1 = 0.92$).

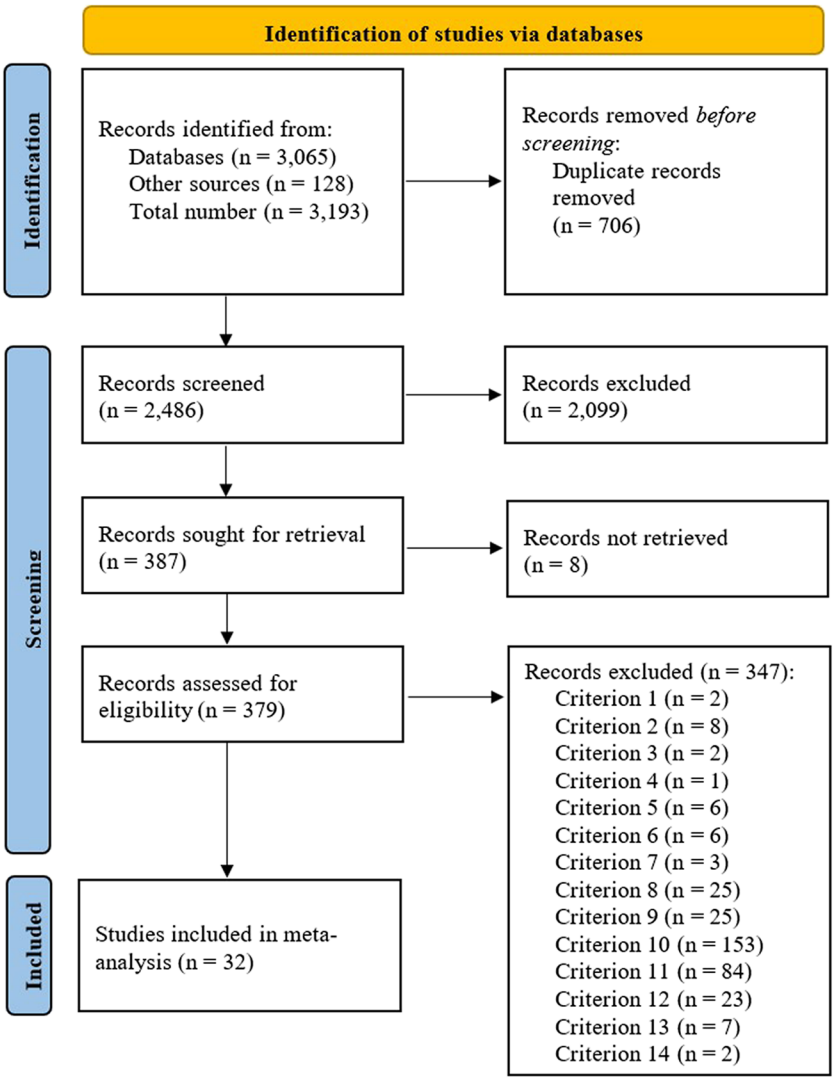

Fig. 2 PRISMA flow diagram of the literature search

Inclusion and Exclusion Criteria

We used several criteria for the inclusion and exclusion of studies. We only included studies published in English (Criterion 1, see Fig. 2) in a peer-reviewed journal (Criterion 2), assuming that papers that have successfully undergone a rigorous peer-review process would provide more reliable results. We excluded studies that did not deal with primary data (e.g., survey articles, reviews; Criterion 3). Furthermore, we excluded all studies that did not actually study peer-feedback (e.g., some studies focus on electromagnetic feedback loops, tutorial feedback, teacher feedback;

Criteria 4 + 5) and studies in which no real peer-feedback took place (e.g., when peer-feedback was provided by fictitious peers; Criterion 6). Next, we excluded all studies that did not take place in a teaching–learning context (e.g., occupational context; Criterion 7). Another important inclusion criterion was that the studies used a (quasi-)experimental design (Criterion 8). With regard to the objects of peer-feedback (according to Alqassab et al., 2023; S. Gielen et al., 2011), we only included studies in which the initial task solution could be assessed separately from the person (Criterion 9; e.g., case analysis, picture, but not an oral presentation or dance performance). We also excluded studies in which the control condition did not receive peer-feedback (Criterion 10). Moreover, to be included, studies had to compare peer-feedback with instructional support vs. peer-feedback without instructional support in at least one condition (Criterion 11). Further, we excluded studies in which effects could not be attributed solely to peer-feedback because peer-feedback was not the only factor that varied across conditions (Criterion 12). Finally, we excluded studies with insufficient statistical reporting, which made it impossible to calculate effect sizes (Criterion 13). We further excluded studies that only reported effect sizes on non-cognitive variables as we are only interested in the effects of instructional support during peer-feedback on cognitive variables (Criterion 14). Decisions on inclusion and exclusion were made by two independent raters with a very good interrater reliability (Gwet's $AC1 = 0.92$). All disagreements were resolved by discussion.

Coding

The 32 studies were coded using a specific category system. This coding scheme was applied to a random subsample of the studies ($N = 9$, $k = 83$) by two independent coders, with sufficient interrater reliability, as described below. Again, all discrepancies were resolved by discussion. The remaining studies were coded by one coder each.

Basic Study Characteristics

First, the basic characteristics of the studies were coded, such as publication year (Gwet's $AC1 = 1.00$) and publication region (USA, Europe, Asia, Oceania; Gwet's $AC1 = 1.00$). In addition, the sample size (Gwet's $AC1 = 1.00$) and the population of the study (school students vs. university students; Gwet's $AC1 = 1.00$) were extracted. To better describe the sample, we further recorded the mean age (Gwet's $AC1 = 0.91$) and standard deviation of age of the participants (Gwet's $AC1 = 0.91$).

Study Design

We further coded whether the study used an experimental or a quasi-experimental design (Gwet's $AC1 = 0.65$), the number of conditions within each study (Gwet's $AC1 = 1.00$), the number of conditions relevant to our meta-analysis (Gwet's $AC1 = 1.00$), and the feedback medium used (analogue vs. digital; Gwet's $AC1 = 0.88$). Note that compared to the other categories, the agreement for the design category

seems rather low. As several studies did not clearly state the procedures they followed to group the students, or used divergent labels in their descriptions, differences between the coders emerged. Still, an acceptable threshold of agreement was reached, and every discrepancy was individually solved through discussion. Similar issues while reviewing studies in the realm of peer-feedback were reported by Alqassab et al. (2023), who reached a comparable interrater agreement. We therefore deem the interrater reliability we achieved as acceptable.

Peer-Feedback Phase Targeted by Instructional Support

For each kind of instructional support that was used in the included studies, we coded the phase during which it was presented. There, we coded whether the instructional support was provided during feedback provision (Gwet's $AC1 = 1.00$) or during feedback reception (Gwet's $AC1 = 0.78$).

Type of Instructional Support

We further coded the instructional support according to its specific focus, distinguishing between preparatory activities, rubrics, sentence starters, guiding questions, and integrated support (Gwet's $AC1 = 0.78$). We formed these categories inductively (e.g., Mayring, 2020) on the basis of the studies identified. After the specific coding of each instructional support during feedback provision and feedback reception, it became clear that each of these types was used in only a few studies. Specifically, only very few studies seem to exist that investigate the effects of these different types of instructional support during feedback reception. Therefore, we decided to create overarching categories (see Online Resource 1) and coded whether the kinds of instructional support had a content-specific or a generic focus (Gwet's $AC1 = 0.79$), based on a distinction that has been repeatedly proposed in instructional research (e.g., Belland et al., 2013; Ertl et al., 2007).

Dependent Variables

We also coded the kinds of dependent variables that were measured in the specific studies. Here, we differentiated between (a) the quality of feedback provision, (b) the quality of feedback reception, and (c) subject-matter-related knowledge (Gwet's $AC1 = 0.82$). Quality of feedback provision is often coded via the quality or quantity of the feedback messages the students provided. There, the quality of the feedback messages can be measured using various criteria (e.g., coding according to correctness and usefulness of feedback; Dmoshinskaia et al., 2021; category system with "peer-feedback style," "verification type," "verification focus," "elaboration type," "elaboration focus"; M. Gielen & De Wever, 2015). Quality of feedback reception is often coded via the quality of the revised task solutions. The main focus here is on the extent to which the feedback received has been implemented and the extent to which the initial task solution has improved as a result (e.g., essay scoring rubric; Altstaedter, 2016; integration of received feedback via comparison of initial and revised task solution; Jurkowski, 2018). As a result, the quality of feedback

reception is often measured with more distal variables than the quality of feedback provision. Subject-matter-related knowledge is typically measured in subsequent post-tests, after the peer-feedback process has been completed (e.g., domain-specific knowledge test; Mulyati & Hadianto, 2023). Thus, while (a) and (b) can be regarded as process variables, (c) can be regarded as an outcome measure.

Effect Sizes Within Studies

In addition to the descriptive characteristics of the studies, the respective effect sizes of the different comparisons within each study were documented (Gwet's $AC1 =$ 0.88 to 0.94). For that sake, we extracted the respective dependent variables with the associated group sizes and effect sizes as means and standard deviations for each condition. This coding forms the basis of the analysis (see Online Resource 1 for an overview of the descriptive statistics of the studies and effect sizes).

Statistical Analyses

We performed all statistical analyses in R [4.3.2], using random effects three-level meta-analysis (e.g., Assink & Wibbelink, 2016; Borenstein et al., 2009). Since many of the included studies measure the same indicators of peer-feedback processes (e.g., quality of feedback provision) or outcomes (e.g., subject-matter-related knowledge) in multiple ways, we used standardized mean differences between the conditions "peer-feedback with instructional support" and "peer-feedback without instructional support" as outcomes (Borenstein et al., 2009; Harrer et al., 2022).[1] With respect to the dependent variables, not all studies ran a comparison between pre-test and post-test data, so we cannot report change scores between the two time points for all studies. Thus, we only used post-test scores in our analyses, following recommendations that change scores and post-test scores should not be mixed, as this would introduce another factor of heterogeneity into the data structure (e.g., Deeks et al., 2019; Lipsey & Wilson, 2001).

To calculate effect sizes (Hedge's g), we used the "metafor" package [4.4–0] (Viechtbauer, 2010). Positive effect sizes indicate higher values in the experimental condition than in the control condition. We used the I^2 statistic to check for heterogeneity of effect sizes (Harrer et al., 2022). The higher the I^2 value, the greater the heterogeneity of effects sizes between studies (Schwarzer et al., 2015). Therefore, this value serves as the basis for calculating moderator analyses with meta-regressions to explain the heterogeneity between studies (Harrer et al., 2022).

We calculated three-level meta-analyses using robust variance estimation (RVE; Hedges et al., 2010) to control for dependencies within studies. Such dependencies include multiple effect sizes for one variable from one study (i.e., when a study included

[1] In one study (i.e., Jurkowski, 2018), the experimental and control conditions consisted of different courses but received exactly the same instructional support or no instructional support at all. Therefore, we combined the scores of the individual courses and calculated an overall score. In other words, we formed an overall score across the different courses for the experimental and control groups.

multiple instruments to measure the same construct) or comparisons of multiple experimental conditions with the same control condition within one study (Hedges et al., 2010; Moeyaert et al., 2017). RVE takes these dependencies within a study into account, without, for example, having to aggregate the individual effect sizes for the same variable within a study (Moeyaert et al., 2017). This means that less information is lost, and statistical analyses can be carried out despite a small number of studies and effect sizes (Moeyaert et al., 2017). We also performed the analyses with RVE with the "metafor" package [4.4–0] (Viechtbauer, 2010) and the "clubSandwich" package [0.5.0] (Pustejovsky, 2023) and used a correlated hierarchical effects structure. Since the correlations within a study are not known, we used rho = 0.6 (Pustejovsky & Tipton, 2022).

To statistically identify potential outliers, we performed influence diagnostics (Cook's Distance and DFBETAS) for all effect sizes (Viechtbauer & Cheung, 2010). To test for publication bias, we used counter-enhanced funnel plots for visual inspection, the rank test as non-parametric version of robustness (Begg & Mazumdar, 1994), and Egger's regression test (Egger et al., 1997) for statistical analysis. Given the way the rank test and Egger's regression test have been developed and validated, they necessitate statistical independence (e.g., Terrin et al., 2003). Consequently, we applied them to a set of independent effect sizes. This means that we aggregated the dependent effect sizes of the individual studies to generate a data set with independent effect sizes (i.e., Borenstein et al., 2009). We chose these methods because they provide a sophisticated insight into the data and can be used in conjunction with RVE.

Results

The results of the influence diagnostics revealed one effect size as outlier (see Online Resource 2). We therefore excluded the outlier from the analyses. Table 1 gives an overview of the included $N = 32$ studies with $N = 3806$ learners. The forest plot with the standardized mean effect sizes for each effect can be found in Online Resource 3.

Publication Bias

The result of Egger's regression test ($b = -0.96$, $t = 2.11$, df = 29, 95% CI [− 1.90, –0.03], $p = 0.04$) was significant, whereas Begg's correlation rank test (Kendall's tau = 0.20, $p = 0.12$) was non-significant. On the one hand, these results indicate the possibility of publication bias. Yet, on the other hand, the heterogeneity of the studies is very high, suggesting that Egger's regression test might be biased. Furthermore, the counter-enhanced funnel plot in Fig. 3 showed a rather symmetrical distribution, indicating no strong evidence for publication bias in our meta-analysis.

Research Question 1: General Effect of Peer-Feedback with Instructional Support vs. Without Instructional Support

As expected, the results regarding the first research question revealed a positive, small-sized general effect of peer-feedback with instructional support compared

Table 1 Description of included studies

Author(s) and year	*N*	*Population*	*Relevant groups*	*Feedback provision support*	*Feedback reception support*	*Dependent variables*
Altstaedter, 2016	70	University students	2	Content-specific	/	Quality of feedback reception
Berg, 1999	46	University students	2	Content-specific	/	Quality of feedback reception
Bouwer et al., 2024[a]	84	School students	2	Content-specific	/	Quality of feedback provision, quality of feedback reception
Bouwer & van der Veen, 2024	212	School students	2	Content-specific	/	Subject-matter-related knowledge
Bürgermeister et al., 2021	385	University students	4	Content-specific, generic	/	Quality of feedback provision
Dmoshinskaia et al., 2021	93	School students	2	Content-specific	/	Quality of feedback provision, quality of feedback reception, subject-matter-related knowledge
Gan & Hattie, 2014	121	School students	2	Generic	/	Quality of feedback provision, quality of feedback reception
M. Gielen & De Wever, 2012	179	University students	2	Generic	/	Quality of feedback provision, quality of feedback reception, subject-matter-related knowledge
M. Gielen & De Wever, 2015	168	University students	3	Generic	/	Quality of feedback provision
S. Gielen et al., 2010	85	School students	2	Content-specific	/	Subject-matter-related knowledge
Gyamfi, et al., 2024	195	University students	2	Generic	/	Quality of feedback provision
Hsu, 2016	224	School students	2	Content-specific	/	Subject-matter-related knowledge
Jurkowski, 2018	70	University students	2	/	Content-specific	Quality of feedback reception
Könings et al., 2019	236	School students	2	Content-specific	/	Quality of feedback provision, quality of feedback reception
Latifi et al., 2021b	52	University students	3	Content-specific, generic	/	Quality of feedback provision, quality of feedback reception, subject-matter-related knowledge

Table 1 (continued)

Author(s) and year	*N*	*Population*	*Relevant groups*	*Feedback provision support*	*Feedback reception support*	*Dependent variables*
Latifi et al., 2021a	86	University students	4	Generic	/	Quality of feedback provision, quality of feedback reception, subject-matter-related knowledge
Lee et al., 2021	40	School students	2	Content-specific	/	Quality of feedback reception, subject-matter-related knowledge
Leenknecht & Prins, 2018	95	School students	2	Content-specific	/	Quality of feedback provision, subject-matter-related knowledge
Mulyati & Hadianto, 2023	270	University students	3	Content-specific	/	Quality of feedback provision, quality of feedback reception, subject-matter-related knowledge
Panadero et al., 2013	209	University students	2	Content-specific	/	Quality of feedback provision
Peters et al., 2018	42	University students	2	Content-specific	/	Quality of feedback reception
Pham et al., 2022	64	School students	2	Generic	/	Subject-matter-related knowledge
Prilop & Weber, 2023	68	University students	2	Content-specific	/	Quality of feedback provision
Sippel, 2024	78	University students	2	Content-specific	/	Subject-matter-related knowledge
Sluijsmans et al., 2004	93	University students	2	Generic	/	Quality of feedback provision, quality of feedback reception, subject-matter-related knowledge
Tsai & Chuang, 2013	48	University students	2	Generic	/	Quality of feedback reception
Valero et al., 2024	221	University students	4	Generic	/	Quality of feedback provision, quality of feedback reception, subject-matter-related knowledge
Wichmann et al., 2018	73	University students	2	/	Generic	Subject-matter-related knowledge
Zeeb et al., 2024[b]	119	University students	4	Content-specific, generic	/	Subject-matter-related knowledge

Table 1 (continued)

Author(s) and year	*N*	*Population*	*Relevant groups*	*Feedback provision support*	*Feedback reception support*	*Dependent variables*
Zhao et al., 2023	41	University students	2	Content-specific	/	Quality of feedback provision, quality of feedback reception
Zheng et al., 2017	64	University students	2	/	Generic	Quality of feedback reception, subject-matter-related knowledge
Zhu, 1995	169	University students	2	Content-specific	/	Quality of feedback provision

[a]Bouwer et al. (2024) analyze a subsample of Bouwer & van der Veen (2024)

[b]Zeeb et al. (2024) analyze a subsample of Bürgermeister et al. (2021). These samples are therefore not included in the total sample

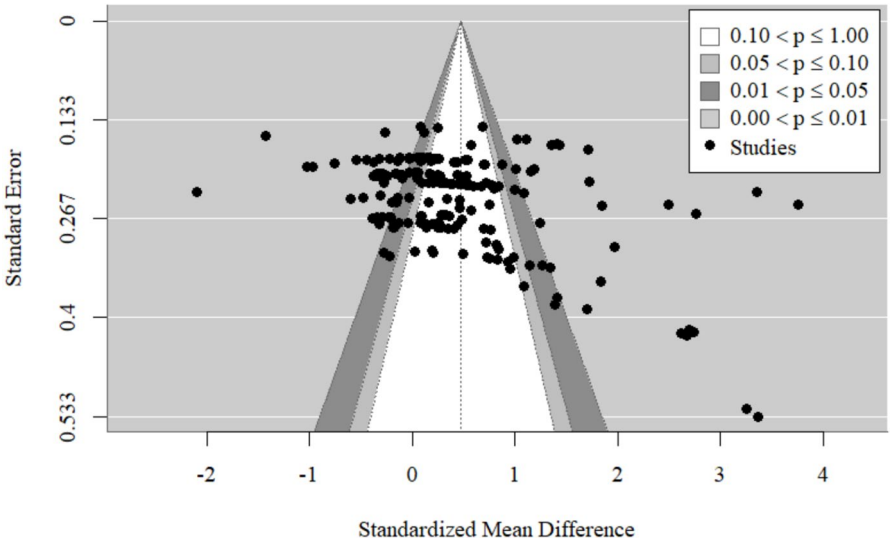

Fig. 3 Counter-enhanced funnel plot to assess publication bias for the effect sizes

to peer-feedback without instructional support ($g = 0.47$, 95% CI [0.23, 0.72], $p < 0.001$). Thus, integrating instructional support in the peer-feedback process seems to promote peer-feedback in general (i.e., without differentiating between different kinds of instructional support and dependent variables), beyond the positive effects of peer-feedback when it is not supported.

Research Question 2: Effects of Peer-Feedback with Instructional Support vs. Without Instructional Support on Different Dependent Variables

Next, we looked at the isolated effects of peer-feedback with vs. without instructional support on (a) the quality of feedback provision (e.g., quality of the feedback messages), (b) the quality of feedback reception (e.g., revision quality), and (c) subject-matter-related knowledge (e.g., domain-specific knowledge). Results revealed a positive, medium-sized effect for peer-feedback with instructional support vs. without instructional support on the quality of feedback provision and a small positive effect on the quality of feedback reception. For subject-matter-related knowledge, we found no significant effect (see Table 2). We also tested whether the effects regarding the different dependent variables were significantly different from each other. This was not the case ($F(2,187) = 1.83$, $p = 0.23$).

Research Question 3: General Effects of Different Kinds of Instructional Support on Peer-Feedback Processes and Outcomes

Regarding the third research question, we compared the general effects of feedback provision support vs. feedback reception support (3a) and of content-specific

Table 2 Overall effect of peer-feedback with instructional support vs. without instructional support on different dependent variables

Dependent variable	*n*	*k*	*g*	*SE*	*dfs*	*95% CI*	*p*
Quality of feedback provision	18	92	0.64	0.16	17.81	[0.30, 0.98]	< 0.01
Quality of feedback reception	18	40	0.49	0.21	20.70	[0.05, 0.92]	< 0.05
Subject-matter-related knowledge	17	60	0.25	0.12	17.31	[− 0.00, 0.51]	0.05

n, number of studies; *k*, number of effect sizes

support vs. generic support *during the specific phases in the peer-feedback process* on peer-feedback processes and outcomes (3b + 3c). For the complete results, see Table 3.

Regarding RQ3a, we found a significantly positive, medium-sized effect of feedback provision support compared to peer-feedback without instructional support. For feedback reception support, we found no significant effect. Yet, as only very few studies exist that investigated the effects of instructional support for feedback reception (as indicated by the low degrees of freedom), this result might not be robust once the evidence basis increases (Fisher & Tipton, 2015). Regarding RQ3b, we found small to medium-sized positive effects for content-specific and generic support during the feedback provision phase compared to peer-feedback without instructional support during the feedback provision phase. Regarding the effects of content-specific support and generic support during the feedback reception phase (RQ3c), we found no significant effects. Again, this might be attributed to the small number of related studies, that is, as the degrees of freedom are all less than four, these results cannot be interpreted (Fisher & Tipton, 2015). We also examined whether the effect sizes of the different types of instructional support on peer-feedback processes and outcomes differed significantly from each other. The results showed no significant differences between the effect sizes for the comparison between feedback provision support and peer-feedback without instructional support on the one hand and the comparison between feedback reception support vs. peer-feedback without

Table 3 General effects of different kinds of instructional support on peer-feedback processes and outcomes

Instructional support	*n*	*k*	*g*	*SE*	*dfs*	*95% CI*	*p*
Feedback provision	29[a]	178	0.51	0.13	25.73	[0.24, 0.78]	< 0.001
Content-specific	20	99	0.48	0.18	14.89	[0.09, 0.87]	< 0.05
Generic	13	79	0.55	0.13	11.94	[0.26, 0.83]	< 0.01
Feedback reception	3	14	0.19	0.23	1.97	[− 0.83, 1.22]	0.50[b]
Content-specific	1	4	0.37	0.13	1	[− 1.32, 2.06]	0.22[b]
Generic	2	10	0.12	0.37	1	[− 4.56, 4.80]	0.80[b]

n, number of studies, *k*, number of effect sizes. [a]The total number of studies of content-specific and generic support differs from the number of studies of feedback provision support, as there are studies that used both content-specific and generic support during feedback provision. [b]As the degrees of freedom are all less than 4, these results cannot be interpreted

instructional support on the other ($F(1, 2.43) = 1.36$, $p = 0.35$). Also, the effects regarding the comparisons between content-specific support during peer-feedback and peer-feedback without instructional support on the one hand and generic support during peer-feedback and peer-feedback without instructional support on the other hand were not significantly different from each other during the feedback provision phase ($F(1, 3.84) = 0.14$, $p = 0.72$) as well as during the feedback reception phase ($F(1, 1.07) = 0.40$, $p = 0.63$).

Research Question 4: Effects of Different Kinds of Instructional Support on the Quality of Feedback Provision, the Quality of Feedback Reception, and Subject-Matter-Related Knowledge

Regarding RQ4, our analyses revealed small to medium-sized positive effects of feedback provision support on the quality of feedback provision and on the quality of feedback reception. For subject-matter-related knowledge, we did not find a significant effect of feedback provision support (see Table 4). To examine the effects of feedback reception support on the different dependent variables, the number of available studies unfortunately was too small to reasonably conduct a meta-analysis. Therefore, these analyses cannot be interpreted. Again, we tested the effect sizes for differences but found no significant differences between the effect sizes of the types of instructional support, neither on the quality of feedback reception ($F(1, 2.87) = 1.74$, $p = 0.28$) nor on subject-matter-related knowledge ($F(1, 1.31) = 0.55$, $p = 0.57$).[2]

Lastly, we were interested in the effects of content-specific and generic support during the specific phases of the peer-feedback process on the quality of feedback provision, the quality of feedback reception, and subject-matter-related knowledge (RQs 4b + 4c).

As Table 4 shows for RQ4b, we found a positive medium-sized effect of content-specific support during feedback provision on the quality of feedback provision but no significant effect on the quality of feedback reception, nor on subject-matter-related knowledge. For generic support during feedback provision, we found positive medium-sized effects on the quality of feedback provision and on subject-matter-related knowledge but no significant effect on the quality of feedback reception. Regarding the differences between the effect sizes, we did not find differences of the types of instructional support for the quality of feedback provision ($F(1, 10.86) = 0.01$, $p = 0.91$), the quality of feedback reception ($F(1, 5.88) = 1.02$, $p = 0.35$), or subject-matter-related knowledge ($F(1, 1.2) = 18.54$, $p = 0.11$).

For RQ4c, we found no significant effects neither of content-specific support nor of generic support during feedback reception on the quality of feedback reception and on subject-matter-related knowledge. Unfortunately, the number of studies available to examine the effects of content-specific support and generic support

[2] Note that we did not examine the effects of feedback reception support on the quality of feedback provision. The reason is that feedback reception support is introduced *after* feedback provision—thus, it is logically not possible for this instructional support to influence feedback provision as well.

Table 4 Effects of different kinds of instructional support on the quality of feedback provision, the quality of feedback reception, and subject-matter-related knowledge

Instructional support	*n*	*k*	*g*	*SE*	*dfs*	*95% CI*	*p*
Dependent variable: Quality of feedback provision							
Feedback provision	18[a]	92	0.72	0.18	13.82	[0.32, 1.12]	< 0.01
Content-specific	13	47	0.75	0.28	8.30	[0.11, 1.38]	< 0.05
Generic	8	45	0.70	0.23	6.61	[0.15, 1.25]	< 0.05
Dependent variable: Quality of feedback reception							
Feedback provision	16[a]	32	0.69	0.26	13.92	[0.12, 1.25]	< 0.05
Content-specific	10	18	0.92	0.42	8.17	[− 0.04, 1.89]	0.06
Generic	7	14	0.40	0.31	5.97	[− 0.36, 1.17]	0.24
Feedback reception	2	8	0.14	0.32	1.98	[− 1.24, 1.53]	0.70[b]
Content-specific	1	4	0.37	0.25	1.00	[− 2.78, 3.52]	0.38[b]
Generic	1	4	0.03	0.52	1.00	[− 6.55, 6.61]	0.96[b]
Dependent variable: Subject-matter-related knowledge							
Feedback provision	15[a]	54	0.43	0.21	12.95	[− 0.03, 0.88]	0.06
Content-specific	10	34	0.34	0.21	12.01	[− 0.12, 0.81]	0.14
Generic	7	20	0.55	0.19	11.02	[0.13, 0.98]	< 0.05
Feedback reception	2	6	0.18	0.26	1.00	[− 3.08, 3.44]	0.60[b]
Content-specific	0						
Generic	2	6	0.17	0.26	1.00	[− 3.12, 3.46]	0.63[b]

n, number of studies; *k*, number of effect sizes. [a]The total number of studies of content-specific and generic support differs from the number of studies of feedback provision support, as there are studies that used both content-specific and generic support during feedback provision. [b]As the degrees of freedom are all less than 4, these results cannot be interpreted

during feedback reception on the quality of feedback reception, and subject-matter-related knowledge (RQ4c) was too small to interpret them as the degrees of freedom are less than four in all these analyses (Fisher & Tipton, 2015).

Discussion

While peer-feedback can be an effective way to promote learning, students often need support regarding how to provide and/or process feedback they receive from their peers to truly benefit from it (e.g., Carless & Boud, 2018; Lui & Andrade, 2022; Strijbos et al., 2010; Winstone et al., 2017). However, prior empirical research (a) has used a broad variety of different kinds of instructional support in peer-feedback scenarios and has yielded (b) rather divergent findings on their effectiveness. Therefore, this meta-analysis tried to uncover the conditions under which instructional support improves the effectiveness of peer-feedback. In doing so, we moved from a general level (comparing the overall effect of peer-feedback with instructional support vs. peer-feedback without instructional support) to more specific levels in order to look more closely at the effects of (a) specific kinds of instructional

support (feedback provision support vs. feedback reception support; content-specific support vs. generic support) that can (b) be presented either during feedback provision or feedback reception on (c) three different kinds of dependent variables, namely, the quality of feedback provision, the quality of feedback reception, and students' subject-matter-related knowledge. The results of our meta-analysis are summarized in Fig. 4.

As we expected for the first research question, we found a significant, medium-sized positive general main effect of peer-feedback with instructional support compared to peer-feedback without instructional support (Li et al., 2020). Thus, we can conclude that peer-feedback can generally be optimized through the integration of instructional support. At first sight, the size of the effect we found may appear a bit disappointing (see Fig. 4 at the top left RQ1). Yet, it has to be noted that our meta-analysis did not compare peer-feedback with instructional support with no-peer-feedback conditions but instead with conditions in which learners also participated in peer-feedback. As peer-feedback in itself has been shown to be an effective learning method (e.g., Double et al., 2020), increasing its effectiveness by almost one half of a standard deviation through the integration of instructional support is definitely notable and strongly suggests that it is worthwhile for practitioners to ponder about including further instructional support in the peer-feedback process. Nevertheless, future meta-analyses should compare peer-feedback with instructional support not only with peer-feedback without instructional support but also with other forms of instructional support. Based on the results of our meta-analysis, we would expect even higher effect sizes in such studies.

To get closer to an understanding of how exactly instructional support during peer-feedback works, we then differentiated between three kinds of dependent

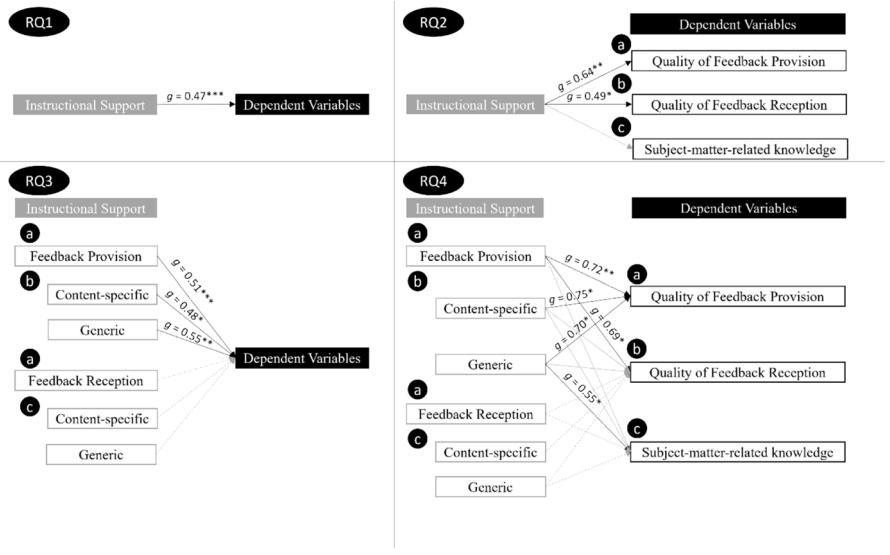

Fig. 4 Graphical representation of all significant effects. Note. Grey arrows indicate non-significant effects and dashed lines indicate non-interpretable results. ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$

variables on which instructional support during peer-feedback might have effects: the quality of feedback provision, the quality of feedback reception, and subject-matter-related knowledge. Our analyses revealed positive effects for the quality of feedback provision and for the quality of feedback reception but with descending effect sizes: While the effect of instructional support on the quality of feedback provision was moderate to large, it was slightly (though not significantly) smaller for the quality of feedback reception. In contrast, we did not find a significant effect for the subject-matter-related knowledge (see Fig. 4 at the top right RQ2). In general, these findings show that augmenting peer-feedback with instructional support yields positive effects on relevant learning processes in peer-feedback (Bürgermeister et al., 2021; Carless & Boud, 2018; Lui & Andrade, 2022). Interestingly though, when taking into account how the quality of feedback provision and feedback reception have been operationalized in the primary studies, it appears that most of them have mainly focused on students' overt behavior (i.e., the way they actually formulated their feedback or the way they edited their initial task solutions based on the feedback they received). Studies that more closely investigate the internal, that is the covert, cognitive processes that learners engage in during feedback provision (i.e., explaining and evaluation processes; van Popta et al., 2017) and feedback reception (i.e., reflect and compare critical the received feedback with their own initial task solution; Nicol et al., 2014), and how they are affected by different kinds of instructional support, are desperately needed. Surprisingly, as the effect of instructional support on subject-matter-related knowledge was not significant, an engagement in these higher-quality cognitive processes does not necessarily translate into the development of a deeper understanding of subject-matter information. This finding is in line with the results of other meta-analyses, which also found no effects in moderator analyses of different types of instructional support on learning outcomes that are assessed after peer-feedback scenarios (e.g., Double et al., 2020; Huisman et al., 2019; Li et al., 2020). Future studies should therefore investigate in more detail how instructional support in the peer-feedback process could be designed to have a positive influence on subject-matter related knowledge.

In addition, for the third research question, we increased the differentiation at the level of the independent variable by asking for the (possibly differential) effects of *different kinds* of instructional support. At the conceptual level, we distinguished between different kinds of instructional support in two ways. First, we were interested in how the general effects of instructional support that is presented during the feedback provision phase would compare with the effects of instructional support that is presented during the feedback reception phase. Our analyses revealed a positive medium-sized effect of feedback provision support but no effect of feedback reception support (see Fig. 4 at the bottom left RQ3a). These effects may indicate that it might be easier to support the feedback provision phase than the feedback reception phase. When providing feedback, the focus of a learner typically is more or less solely on their peer's initial performance, as they evaluate someone else's rather than their own task solution. In contrast, the phase of receiving feedback appears to be more complex (e.g., Lipnevich & Smith, 2022; Winstone et al., 2017): For feedback to be used efficiently, the feedback receiver must understand it in the first place. Second, even when students

have understood the feedback, their own perceptions and skills play a crucial role in how they use the feedback they receive (Berndt et al., 2018; Strijbos et al., 2021). Thus, receiving feedback involves both processing and implementing the feedback received, which may require different and more fine-grained types of instructional support than providing feedback (e.g., Li & Grion, 2019). However, when we look at the number of studies and the effect sizes of feedback reception support, it is striking that only very few studies address the promotion of feedback reception at all. Therefore, the lack of effects of feedback reception support may also be caused by a statistical power problem (Fisher & Tipton, 2015). Definitely more research is necessary, especially on the effects of instructional support that is presented during feedback reception.

An alternative differentiation between different kinds of instructional support we made was between content-specific and generic support during the specific phases of the peer-feedback process. Content-specific support refers to the provision of content-related support for a specific learning objective, while generic support refers to structuring the peer-feedback process without a specific content-related reference. Here, we found a significant small to medium-sized positive effect for content-specific support and generic support during feedback provision but not during feedback reception (see Fig. 4 at the bottom left RQ3b). As there is no significant difference between the effect of content-specific support and generic support during feedback provision, it seems that, in principle, both types of instructional support can be used during feedback provision. For practitioners, it might be easier then to use generic support, as it can be applied across different contexts and domains (Hein & Prediger, 2017; Reiser, 2004). Content-specific support, in contrast, needs to be tailored to each new learning situation, which makes it costly to develop (e.g., Hannafin et al., 1999; Vogel et al., 2017).

The fact that we did not find effects of content-specific and generic support on feedback reception may again indicate that feedback reception is more difficult to support due to the reasons outlined above (see Fig. 4 at the bottom left RQ3c). Yet again, the small number of available studies may also hint towards a statistical power problem (Fisher & Tipton, 2015). Since the processes enfolding in the feedback reception phase are very complex, and since students often have difficulties processing the feedback they receive (e.g., Winstone et al., 2017), especially if they have issues understanding the feedback (e.g., Wu & Schunn, 2020a), it is possible that content-specific support would be an efficient way of providing support, as it would assist students to elaborate the learning content more deeply. This, in turn, might help them understand the content and the respective feedback more deeply. Another hypothesis might be that the main problem is the integration of multiple feedback messages rather than the implementation of a single feedback. The few studies that exist do not address this integration. These hypotheses should be tested in future studies.

Finally, the fourth research question asked for possible interactions of different kinds of instructional support and the three different dependent variables we investigated (quality of feedback provision, quality of feedback reception, and subject-matter-related knowledge). Our analyses revealed positive, medium-sized effects of feedback provision support on the quality of feedback provision and on

the quality of feedback reception but not on subject-matter-related knowledge. In addition, we found no significant effects of feedback reception support on the quality of feedback reception nor on subject-matter-related knowledge (see Fig. 4 at the bottom right RQ4a). The results regarding feedback provision imply that supporting students with respect to how to provide high quality feedback messages does not only lead to improvements in this very phase; it also has spillover effects to the feedback reception phase. In other words, if learners process the subject-matter better during feedback provision through instructional support, this could lead to better skill acquisition and thus to better revision performance in the feedback reception phase (e.g., Wu & Schunn, 2020b). We would also have expected feedback provision support to have a positive effect on subject-matter-related knowledge, as a deeper elaboration of content material as well as of a peer's feedback should ultimately also lead to greater subject-matter-related knowledge (e.g., Ion et al., 2018). However, we did not find this in the meta-analysis. Thus, besides the fact that knowledge post-tests represent rather distal measures of the effects of peer-feedback (see above), there may be further processes that occur during peer-feedback which might be more predictive for subject-matter-related knowledge that are typically not measured in related studies. Given our observation that most studies we included strongly focused on overt processes that occur during peer-feedback, an engagement in covert, subject-matter related cognitive processes might be more responsible for subject-matter-related knowledge (e.g., Stegmann et al., 2012). Further research is necessary to confirm or disconfirm this assumption.

From a theoretical point of view, we would also have expected positive effects of instructional support that specifically targets the feedback reception phase, as also during this phase meaningful learning processes such as critical thinking should take place (e.g., Winstone et al., 2017). One explanation for the lack of such effects might be that only very few studies measure learning processes that capture how learners actually *process* peer-feedback. For example, in a study by M. Gielen and De Wever (2012), the most proximal variable that was measured regarding feedback reception was the extent to which students used the feedback they received in their revisions. However, as the process model by Bauer et al. (2023) suggests, before actually implementing feedback in their task solution, learners first need to process the feedback and reflect upon it. Therefore, it is not guaranteed that the studies ensured that the students actually engaged in mindful processing, as a kind of manipulation check (Bolzer et al., 2015). If this did not happen, then we do not know whether the reception scaffolds had the chance to have any effect at all. Thus, future studies should develop and use methods to better capture the feedback reception and the covert cognitive processes that occur during this phase (e.g., through think-aloud methods, Stegmann et al., 2012; or via eye-tracking, e.g., Cutumisu et al., 2019). Bolzer et al. (2015), who explored how mindful cognitive processing can be measured via eye-tracking, provide a good starting point for this.

Regarding the second distinction we made with respect to different kinds of instructional support, we found a positive effect of content-specific support on the quality of feedback provision but neither on the quality of feedback reception nor on subject-matter-related knowledge. In addition, we found a significant positive

effect of generic support on the quality of feedback provision and on subject-matter-related knowledge but not on the quality of feedback reception (see Fig. 4 at the bottom right RQ4b). Again, both kinds of support have positive effects on the quality of feedback provision is good news for educators, as any choice they make between content-specific or generic support will likely yield positive effects (e.g., Bouwer & van der Veen, 2024; M. Gielen & De Wever, 2012). Yet, when they choose generic support, they may additionally produce positive effects on students' subject-matter-related knowledge, which according to our results is less likely when they opt for content-specific support (e.g., Hsu, 2016). The significant effect of generic but not of content-specific support during feedback provision on subject-matter-related knowledge is interesting. It is a surprising finding, as we expected that content-specific support in particular would have a positive effect on subject-matter-related knowledge, as this kind of support is primarily concerned with having learners elaborate more deeply on content information (Double et al., 2020; Ion et al., 2018). This, however, does not seem to happen automatically. Apparently, content-specific support needs to be very carefully designed to yield such positive effects. In turn, a possible explanation for the positive effect of generic support on subject-matter-related knowledge could be that this structural support might reduce students' need to overly ponder about how to provide feedback (as generic support will explicitly tell them how to do this), so that they can focus explicitly on the learning content (e.g., Hein & Prediger, 2017; Reiser, 2004). Further research on this topic could be worthwhile. In addition, it is surprising that we did not find an effect of content-specific support during feedback provision on the quality of feedback reception. One explanation for this could be that students have a double task when they are not supported by generic support: on the one hand, they have to follow the content-specific support in order to get through the learning content well, and on the other hand, they have to write well-structured feedback at the same time. This can lead to cognitive overload, so that the potential of content-specific support is lost (Sweller, 1999).

With regard to the effects of content-specific support and generic support during feedback reception on the quality of feedback reception and subject-matter-related knowledge, we found no significant effects (see Fig. 4 at the bottom right RQ4c). However, the evidence base is too weak to arrive at definitive conclusions (Fisher & Tipton, 2015). Nevertheless, it is striking that we did not find any effects of content-specific support on subject-matter-related knowledge in the studies on supporting feedback reception. There is an urgent need for more studies that deal explicitly with how to support feedback reception effectively.

Taken together, the results of this meta-analysis indicate that the general effectiveness of peer-feedback as instructional method can further be improved by adding instructional support to the peer-feedback process, especially by implementing instructional support that supports learners in providing high-quality feedback to each other. Future research should however more thoroughly examine under what circumstances feedback reception support can be beneficial for learning.

Limitations

Of course, our meta-analysis comes with limitations. First, we only included studies that were published in a journal. Of course, there may be other unpublished data or conference papers on peer-feedback with instructional support vs. peer-feedback without instructional support that are not included here. However, we can assume that our results are not biased by this as there is little evidence of publication bias. Future research could also examine unpublished data on this topic.

Second, applying rather strict exclusion criteria allowed us to include only 32 studies, indicating that the number of (quasi-)experimental studies that compare the effects of peer-feedback with instructional support vs. peer-feedback without instructional support is still rather low. In addition, our literature search found only $n = 3$ studies with $k = 14$ effect sizes for feedback reception support at all. This means that the results cannot be interpreted meta-analytically, as the number of effect sizes for individual combinations is far too small. Therefore, in general, further studies that investigate the effects of several kinds of instructional support on peer-feedback processes and subject-matter-related knowledge are desperately needed. However, this notion suggests that more research is clearly needed on effective feedback reception support. Individual studies may focus on specific combinations of support methods and respective fitting dependent variables, such as the effect of content-specific support during feedback reception on the quality of feedback reception. This research is necessary to further improve the effectiveness of peer-feedback processes.

Third, the small N of available studies in general allowed us to look at the types of instructional support and indicators of peer-feedback learning processes and outcomes only at a rather coarse-grained level, as it was only possible to differentiate between feedback provision support and feedback reception support on the one hand, and content-specific support and generic support during the specific phases in the peer-feedback process on the other. It would certainly be interesting to apply more fine-grained differentiations, for example, between instructional support that comes in the form of preparatory activities, rubrics, sentence starters, guiding questions, or integrated (combined) support. Likewise, also the dependent variables could be further subdivided into the quality of feedback messages as part of the quality of feedback provision, revision quality as part of the quality of feedback reception, or learning achievement in a specific context as part of the subject-matter-related knowledge. It would then be interesting to examine the effect of instructional support on these individual dependent variables. As of now, this is not possible based on the current study situation.

Implications

Despite these limitations, our meta-analysis yields important implications for research and practice. With respect to research, it shows that more studies are needed that take a particular look at the feedback reception phase and how it can be structured. Moreover, the results of the meta-analysis offer great potential for theory development. The results show that different types of instructional support

Table 5 Evidence-based practice recommendations on how to support learners during feedback provision

Dependent variables	*Instructional support*	*Example*
Quality of feedback provision	Content-specific	Concepts: (4) All the important and secondary concepts are included. (3) Contains the important and some secondary concepts but not all. (2) The important concepts are included but not the secondary ones. (1) Some key concepts are lacking. (Panadero et al., 2013)
	Generic	"The aim is to…"; "You've already done very well…"; "For the future, hold on to…" (Bürgermeister et al., 2021)
Quality of feedback reception	Content-specific	"Is the operation chosen correctly?", "Why is the operation incorrect?", "How can the working plan be improved?" (Peters et al., 2018)
	Generic	What did he/she do well? Give explanations to support your feedback What didn't he/she do well? Give explanations to support your feedback How can he/she improve on the current piece of work? Give explanations to support your feedback. (Gan & Hattie, 2014)
Subject-matter-related knowledge	Content-specific	Assessment criteria for the criteria condition (translated from Russian) 1. What important concepts are missing? 2. How would you change the structure of the map? 3. Which links should be renamed to be more meaningful? 4. What examples should be added? 5. Why is this concept map helpful or not helpful for understanding the topic? (Dmoshinskaia et al., 2021)
	Generic	The clear position on the topic: Feed Back: To what extent your learning partner provide his/her clear position on the topic? Please explain Feed Forward: What is your advice to your learning partner to (better) provide his/her clear opinion on the topic? Please explain Feed Back + Feed Forward: To what extent your learning partner provide his/her clear position on the topic? What are your suggestions? (Latifi et al., 2021a)

in the two phases of feedback provision and feedback reception can have similar effects on the dependent variables analyzed (i.e., quality of feedback provision, quality of feedback reception, subject-matter-related knowledge). This suggests that there are complex mechanisms at work in the peer-feedback process, which should be uncovered in future research. This is also supported by the fact that the dependent variables are rather coarse-grained and relate to overt rather than covert cognitive processes. In order to be able to better categorize the effects, it would generally be worthwhile to have more studies that explicitly refer to the promotion of specific cognitive processes within the phases of the peer-feedback process. It would be worthwhile to explore this in more detail in the future in order to develop a theoretical model that explicitly addresses the effects of instructional support on different kinds of peer-feedback learning processes and outcomes.

For educational practice, the added value of this meta-analysis is that while peer-feedback by itself is an effective method to support student learning (e.g., Double et al., 2020), it can be even more effective if instructors carefully integrate instructional support in the peer-feedback process. The finding that there are differential effects regarding the interaction of different types of instructional support on different kinds of peer-feedback learning processes and outcomes suggests that instructors should align the instructional support used with the particular peer-feedback learning processes or outcomes they intend to facilitate (for an overview based on our results, see Table 5). For example, based on our findings, when teachers aim to improve the quality of feedback provision, they could either use content-specific support that helps discuss and reflect upon the domain at hand at a deeper level or generic support that structures feedback provision by providing specific instructions on how the feedback should be formulated (e.g., according to feed up, feed back, feed forward; Hattie & Timperley, 2007). Unfortunately though, the overview in Table 5 currently only includes combinations of instructional support and dependent variables for feedback provision, as the current evidence base for feedback reception does not allow for reliable results.

Conclusion

Overall, the results of our meta-analysis show that instructional support, as opposed to no instructional support, can effectively improve the peer-feedback process. There are different types of instructional support that can be related either to the specific phases of the peer-feedback process (feedback provision vs. feedback reception) or to the design of the support within the phases (content-specific vs. generic). In general, these types of instructional support seem to have positive effects on different learning processes (i.e., formulating high-quality feedback messages, effectively elaborating on the feedback received) and learning outcomes (i.e., subject-matter-related knowledge). While we can say this rather reliably with respect to feedback provision, more research is needed at this stage that looks at the effects of instructional support on feedback reception,

which is an important phase in the peer-feedback process. On the other hand, it also shows that more sound, (quasi-)experimental research is needed on specific support interventions, which may further inform strongly needed theory development in this area. Nevertheless, our meta-analysis provides a useful overview of certain combinations of instructional support during feedback provision and dependent variables that can be used for practical implementation of peer-feedback processes.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10648-025-10017-3.

Author Contribution Conceptualization: Julia Hornstein, Melanie V. Keller, Martin Greisel, Markus Dresel, and Ingo Kollar; methodology: Julia Hornstein, Melanie V. Keller, Martin Greisel, Markus Dresel, and Ingo Kollar; formal analysis and investigation: Julia Hornstein and Melanie V. Keller; writing—original draft preparation: Julia Hornstein; writing—review and editing: Melanie V. Keller, Martin Greisel, Markus Dresel, and Ingo Kollar; funding acquisition: Ingo Kollar; supervision: Ingo Kollar.

Funding Open Access funding enabled and organized by Projekt DEAL. This work was supported by the Stiftung Innovation in der Hochschullehre (project "Facilitating Competence Development through Authentic, Digital, and Feedback-Based Teaching–Learning Scenarios", grant number FBM2020).

Data Availability The data is available via the open science framework (https://osf.io/bfj3z/?view_only=be047b4469c74ae08609827e599beab0).

Declarations

Conflict of Interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

References marked with an * were included in the meta-analysis

Alemdag, E., & Yildirim, Z. (2022). Effectiveness of online regulation scaffolds on peer feedback provision and uptake: A mixed methods study. *Computers & Education, 188*, 104574. https://doi.org/10.1016/j.compedu.2022.104574

*Altstaedter, L. L. (2016). Investigating the impact of peer feedback in foreign language writing. *Innovation in Language Learning and Teaching*, *12*(2), 137–151. https://doi.org/10.1080/17501229.2015.1115052

Alqassab, M., Strijbos, J.-W., Panadero, E., Ruiz, J. F., Warrens, M., & To, J. (2023). A systematic review of peer assessment design elements. *Educational Psychology Review, 35*(18), 1–36. https://doi.org/10.1007/s10648-023-09723-7

Assink, M., & Wibbelink, C. J. M. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology, 12*(3), 154–174. https://doi.org/10.20982/tqmp.12.3.p154

Bauer, E., Greisel, M., Kuznetsov, I., Berndt, M., Kollar, I., Dresel, M., Fischer, M. R., & Fischer, F. (2023). Using natural language processing to support peer-feedback in the age of artificial intelligence: A cross-disciplinary framework and a research agenda. *British Journal of Educational Technology, 54*(5), 1222–1245. https://doi.org/10.1111/bjet.13336

Belland, B. R., Gu, J., Armbrust, S., & Cook, B. (2013). Using generic and context-specific scaffolding to support authentic science inquiry. In D. G. Sampson, J. M. Spector, D. Ifenthaler, P. Isaías (Eds.), *Proceeings of IADIS International Conference Cognition and Exploratory Learning in Digital Age - CELDA 2013* (pp. 185–192). International Association for Development of the Information Society, Texas.

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics, 50*(4), 1088–1101. https://doi.org/10.2307/2533446

*Berg, E. (1999). The effects of trained peer response on ESL University students' revision types and writing quality. *Journal of Second Language Writing, 8*(3), 215–241. https://doi.org/10.1016/S1060-3743(99)80115-5

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. https://doi.org/10.1007/s10212-017-0343-z

Bisra, K., Liu, Q., Nesbit, J. C., Salimi, F., & Winne, P. H. (2018). Inducing self-explanation: A meta-analysis. *Educational Psychology Review, 30*(3), 703–725. https://doi.org/10.1007/s10648-018-9434-x

Bolzer, M., Strijbos, J.-W., & Fischer, F. (2015). Inferring mindful cognitive-processing of peer-feedback via eye-tracking: role of feedback-characteristics fixation-durations and transitions. *Journal of Computer Assisted Learning, 31*(5), 422–434. https://doi.org/10.1111/jcal.12091

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley. https://doi.org/10.1002/9780470743386

*Bouwer, R., & van der Veen, C. (2024). Write, talk and rewrite: The effectiveness of a dialogic writing intervention in upper elementary education. *Reading and Writing, 37*(6), 1435–1456. https://doi.org/10.1007/s11145-023-10474-8

*Bouwer, R., van Braak, M., & van der Veen, C. (2024). Dialogic writing in the upper grades of primary school: How to support peer feedback conversations that promote meaningful revisions. *Learning and Instruction, 93*, 101965. https://doi.org/10.1016/j.learninstruc.2024.101965

*Bürgermeister, A., Glogger-Frey, I., & Saalbach, H. (2021). Supporting peer feedback on learning strategies: Effects on self-efficacy and feedback quality. *Psychology Learning & Teaching*, *20*(3), 383–404. https://doi.org/10.1177/14757257211016604

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Çevik, Y. D. (2015). Assessor or assessee? Investigating the differential effects of online peer assessment roles in the development of students' problem-solving skills. *Computers in Human Behavior, 52*, 250–258. https://doi.org/10.1016/j.chb.2015.05.056

Cutumisu, M., Turgeon, K.-L., Saiyera, T., Chuong, S., González Esparza, L. M., MacDonald, R., & Kokhan, V. (2019). Eye tracking the feedback assigned to undergraduate students in a digital assessment game. *Frontiers in Psychology, 10*, 1931. https://doi.org/10.3389/fpsyg.2019.01931

Davis, E. A., & Miyake, N. (2004). Explorations of scaffolding in complex classroom systems. *Journal of the Learning Sciences, 13*(3), 265–272. https://doi.org/10.1207/s15327809jls1303_1

Deeks, J. J., Higgins, J. P. T., & Altman, D. G. (2019). Analysing data and undertaking meta-analyses. In J. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch (Eds.), *Cochrane book series. Cochrane handbook for systematic reviews of interventions* (2nd ed., pp. 241–284). Wiley Blackwell. https://doi.org/10.1002/9781119536604.ch10

*Dmoshinskaia, N., Gijlers, H., & Jong, T. de (2021). Giving feedback on peers' concept maps in an inquiry learning context: The effect of providing assessment criteria. *Journal of Science Education and Technology*, *30*(3), 420–430. https://doi.org/10.1007/s10956-020-09884-y

Double, K. S., McGrane, J. A., & Hopfenbeck, T. N. (2020). The impact of peer assessment on academic performance: A meta-analysis of control group studies. *Educational Psychology Review, 32*(2), 481–509. https://doi.org/10.1007/s10648-019-09510-3

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ (Clinical Research Ed.), 315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Ertl, B., Kopp, B., & Mandl, H. (2007). Supporting collaborative learning in videoconferencing using collaboration scripts and content schemes. In F. Fischer, I. Kollar, H. Mandl, & J. M. Haake (Eds.), *Computer-supported collaborative learning series: Vol. 6. Scripting computer-supported collaborative learning: Cognitive, computational and educational perspectives* (pp. 213–236). Springer Science Business Media LLC. https://doi.org/10.1007/978-0-387-36949-5_13

Falchikov, N. (2013). *Improving assessment through student involvement: Practical solutions for aiding learning in higher and further education*. Routledge. https://doi.org/10.4324/9780203220993

Fisher, Z., & Tipton, E. (2015). *robumeta: Robust variance meta-regression (R package version 2.1)*. https://doi.org/10.48550/arXiv.1503.02220

*Gan, M. J., S, & Hattie, J. (2014).Prompting secondary university students' use of criteria, feedback specificity and feedback levels during an investigative task. *Instructional Science*, *42*(6), 861–878. https://doi.org/10.1007/s11251-014-9319-4

Gielen, S., Dochy, F., & Onghena, P. (2011). An inventory of peer assessment diversity. *Assessment & Evaluation in Higher Education, 36*(2), 137–155. https://doi.org/10.1080/02602930903221444

*Gielen, M., & De Wever, B. (2012). Peer assessment in a wiki: Product improvement, students' learning and perception regarding peer feedback. *Procedia: Social and Behavioral Sciences*, *69*, 585–594. https://doi.org/10.1016/j.sbspro.2012.11.450

*Gielen, M., & De Wever, B. (2015). Structuring peer assessment: Comparing the impact of the degree of structure on peer feedback content. *Computers in Human Behavior*, *52*, 315–325. https://doi.org/10.1016/j.chb.2015.06.019

*Gielen, S., Tops, L., Dochy, F., Onghena, P., & Smeets, S. (2010). A comparative study of peer and teacher feedback and of various peer feedback forms in a secondary school writing curriculum. *British Educational Research Journal*, *36*(1), 143–162. http://www.jstor.org/stable/27823591

*Gyamfi, G., Hanna, B. E., Khosravi, H. (2024). Impact of an instructional guide and examples on the quality of feedback: Insights from a randomised controlled study. *Education Technology Research and Development, 72*(3), 1419–1437. https://doi.org/10.1007/s11423-024-10346-0

Hannafin, M., Land, S., & Oliver, K. (1999). Open learning environments: Foundations, methods, and models. In C. M. Reigeluth (Ed.), *Instructional-design theories and models* (pp. 115–140). Routledge Taylor and Francis Group. https://doi.org/10.4324/9781410603784-8

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2022). *Doing meta-analysis with R: A hands-on guide* (First edition). *Taylor & Francis eBooks*. CRC Press Taylor & Francis Group. https://doi.org/10.1201/9781003107347

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

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1*(1), 39–65. https://doi.org/10.1002/jrsm.5

Hein, K., & Prediger, S. (2017). Fostering and investigating university students' pathways to formal reasoning: A design research project on structural scaffolding for 9th graders. In T. Dooley & G. Gueudet (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education (CERME 10)*. Dublin, Ireland: DCU / ERME, (pp. 163–170).

*Hsu, T.-C. (2016). Effects of a peer assessment system based on a grid-based knowledge classification approach on computer skills training. *Educational Technology and Society*, *19*(4), 100–111.

Huisman, B., Saab, N., van den Broek, P., & van Driel, J. (2019). The impact of formative peer feedback on higher education university students' academic writing: A meta-analysis. *Assessment & Evaluation in Higher Education, 44*(6), 863–880. https://doi.org/10.1080/02602938.2018.1545896

Ion, G., Sánchez Martí, A., & Agud Morell, I. (2018). Giving or receiving feedback: Which is more beneficial to university students' learning? *Assessment & Evaluation in Higher Education, 44*(1), 124–138. https://doi.org/10.1080/02602938.2018.1484881

Jackson, S. L., Stratford, S. J., Krajcik, J., & Soloway, E. (1994). Making dynamic modeling accessible to precollege science students. *Interactive Learning Environments, 4*(3), 233–257. https://doi.org/10.1080/1049482940040305

*Jurkowski, S. (2018). Do question prompts support university students in working with peer feedback? *International Journal of Educational Research*, *92*, 1–9. https://doi.org/10.1016/j.ijer.2018.07.003

Kollar, I., & Fischer, F. (2010). Peer assessment as collaborative learning: A cognitive perspective. *Learning and Instruction, 20*(4), 344–348. https://doi.org/10.1016/j.learninstruc.2009.08.005

Könings, K. D., van Zundert, M., & van Merriënboer, J. J. (2019). Scaffolding peer-assessment skills: Risk of interference with learning domain-specific skills? Learning and Instruction, 60*, 85–94. https://doi.org/10.1016/j.learninstruc.2018.11.007

Lachner, A., Hoogerheide, V., van Gog, T., & Renkl, A. (2022). Learning-by-teaching without audience presence or interaction: When and why does it work? *Educational Psychology Review, 34*(2), 575–607. https://doi.org/10.1007/s10648-021-09643-4

*Latifi, S., Noroozi, O., & Talaee, E. (2021a). Peer feedback or peer feedforward? Enhancing students' argumentative peer learning processes and outcomes. *British Journal of Educational Technology*, *52*(2), 768–784. https://doi.org/10.1111/bjet.13054

*Latifi, S., Noroozi, O., Hatami, J., & Biemans, H. J. (2021b). How does online peer feedback improve argumentative essay writing and learning? *Innovations in Education and Teaching International*, *58*(2), 195–206. https://doi.org/10.1080/14703297.2019.1687005

*Lee, Y. F., Lin, C. J., Hwang, G. J., Fu, Q. K., & Tseng, W. H. (2021). Effects of a mobile-based progressive peer-feedback scaffolding strategy on students' creative thinking performance, metacognitive awareness, and learning attitude. *Interactive Learning Environments*, *31*(5), 2986–3002. https://doi.org/10.1080/10494820.2021.1916763

*Leenknecht, M. J., & Prins, F. J. (2018). Formative peer assessment in primary school: The effects of involving pupils in setting assessment criteria on their appraisal and feedback style. *European Journal of Psychology of Education*, *33*(1), 101–116. https://doi.org/10.1007/s10212-017-0340-2

Li, L., & Grion, V. (2019). The power of giving feedback and receiving feedback in peer assessment. *AISHE-J, 11*(2), 1–17. https://doi.org/10.62707/aishej.v11i2.413

Li, H., Xiong, Y., Hunter, C. V., Guo, X., & Tywoniw, R. (2020). Does peer assessment promote student learning? A meta-analysis. *Assessment & Evaluation in Higher Education, 45*(2), 193–211. https://doi.org/10.1080/02602938.2019.1620679

Lipnevich, A., & Smith, J. K. (2022). Student – Feedback interaction model: Revised. *Studies in Educational Evaluation, 75*, 101208. https://doi.org/10.1016/j.stueduc.2022.101208

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

Lui, A. M., & Andrade, H. L. (2022). The next black box of formative assessment: A model of the internal mechanisms of feedback processing. *Frontiers in Education, 7*, 751548. https://doi.org/10.3389/feduc.2022.751548

Lundstrom, K., & Baker, W. (2009). To give is better than to receive: The benefits of peer review to the reviewer's own writing. *Journal of Second Language Writing, 18*(1), 30–43. https://doi.org/10.1016/j.jslw.2008.06.002

Mayring, P. (2020). Qualitative Inhaltsanalyse [Qualitative Content Analysis]. In G. Mey & K. Mruck (Eds.), *Handbuch Qualitative Forschung in der Psychologie: Band 2: Designs und Verfahren* (2., erw. u. überarb. Auflage 2020, pp. 495–511) [*Handbook Qualitative Research in Psychology: Volume 2: Designs and Procedures* (2., extended and revised edition 2020, pp. 495–511]. Springer Fachmedien Wiesbaden GmbH. https://doi.org/10.1007/978-3-658-26887-9_52

Moeyaert, M., Ugille, M., Beretvas, S. N., Ferron, J., Bunuan, R., & van den Noortgate, W. (2017). Methods for dealing with multiple outcomes in meta-analysis: A comparison between averaging effect sizes, robust variance estimation and multilevel meta-analysis. *International Journal of Social Research Methodology, 20*(6), 559–572. https://doi.org/10.1080/13645579.2016.1252189

*Mulyati, Y., & Hadianto, D. (2023). Enhancing argumentative writing via online peer feedback-based essay: A quasi-experiment study. *International Journal of Instruction, 16*(2), 195–212. https://doi.org/10.29333/iji.2023.16212a

Nicol, D., & McCallum, S. (2022). Making internal feedback explicit: Exploiting the multiple comparisons that occur during peer review. *Assessment & Evaluation in Higher Education, 47*(3), 424–443. https://doi.org/10.1080/02602938.2021.1924620

Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: A peer review perspective. *Assessment & Evaluation in Higher Education, 39*(1), 102–122. https://doi.org/10.1080/02602938.2013.795518

*Panadero, E., Romero, M., & Strijbos, J.-W. (2013). The impact of a rubric and friendship on peer assessment: Effects on construct validity, performance, and perceptions of fairness and comfort. *Studies in Educational Evaluation*, *39*(4), 195–203. https://doi.org/10.1016/j.stueduc.2013.10.005

Patchan, M. M., & Schunn, C. D. (2015). Understanding the benefits of providing peer feedback: How students respond to peers' texts of varying quality. *Instructional Science, 43*(5), 591–614. https://doi.org/10.1007/s11251-015-9353-x

*Peters, O., Körndle, H., & Narciss, S. (2018). Effects of a formative assessment script on how vocational students generate formative feedback to a peer's or their own performance. *European Journal of Psychology of Education*, *33*(1), 117–143. https://doi.org/10.1007/s10212-017-0344-y

*Pham, V. P. H. P., Le, T. C., Phan, T. H., & Nguyen, N. H. V. (2022). The effects of trained peer feedback for high school students. *World Journal of English Language*, *12*(1), 27. https://doi.org/10.5430/wjel.v12n1p27

*Prilop, C. N., & Weber, K. E. (2023). Digital video-based peer feedback training: The effect of expert feedback on pre-service teachers' peer feedback beliefs and peer feedback quality. *Teaching and Teacher Education, 127*, 104099. https://doi.org/10.1016/j.tate.2023.104099

Pustejovsky, J. (2023). clubSandwich: Cluster-Robust (Sandwich) Variance Estimators with Small-Sample Corrections. R package version 0.5.10, https://CRAN.R-project.org/package=clubSandwich

Pustejovsky, J. E., & Tipton, E. (2022). Meta-analysis with robust variance estimation: Expanding the range of working models. *Prevention Science, 23*(3), 425–438. https://doi.org/10.1007/s11121-021-01246-3

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *Journal of the Learning Sciences, 13*(3), 273–304. https://doi.org/10.1207/s15327809jls1303_2

Schwarzer, G., Carpenter, J. R., & Rücker, G. (2015). *Meta-analysis with R*. Springer. https://doi.org/10.1007/978-3-319-21416-0

Simons, K. D., & Klein, J. D. (2007). The impact of scaffolding and student achievement levels in a problem-based learning environment. *Instructional Science, 35*(1), 41–72. https://doi.org/10.1007/s11251-006-9002-5

*Sippel, L. (2024). Maximizing the benefits of peer interaction: Form-focused instruction and peer feedback training. *Language Teaching Research, 28*(2), 413–439. https://doi.org/10.1177/13621688211004638

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

Stegmann, K., Wecker, C., Weinberger, A., & Fischer, F. (2012). Collaborative argumentation and cognitive elaboration in a computer-supported collaborative learning environment. *Instructional Science, 40*(2), 297–323. https://doi.org/10.1007/s11251-011-9174-5

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. https://doi.org/10.1016/j.learninstruc.2009.08.008

Strijbos, J.-W., Pat-El, R., & Narciss, S. (2021). Structural validity and invariance of the feedback perceptions questionnaire. *Studies in Educational Evaluation, 68*, 100980. https://doi.org/10.1016/j.stueduc.2021.100980

Sweller, J. (1999). *Instructional design in technical areas*. ACER Press.

Terrin, N., Schmid, C. H., Lau, J., & Olkin, I. (2003). Adjusting for publication bias in the presence of heterogeneity. *Statistics in Medicine, 22*(13), 2113–2126. https://doi.org/10.1002/sim.1461

*Tsai, Y.-C., & Chuang, M.-T. (2013). Fostering revision of argumentative writing through structured peer assessment.*Perceptual and Motor Skills*, *116*(1), 210–221. https://doi.org/10.2466/10.23.PMS.116.1.210-221

Valero Haro, A., Noroozi, O., Biemans, H. J. A., & Mulder, M. (2019). The effects of an online learning environment with worked examples and peer feedback on students' argumentative essay writing and domain-specific knowledge acquisition in the field of biotechnology. *Journal of Biological Education, 53*(4), 390–398. https://doi.org/10.1080/00219266.2018.1472132

*Valero Haro, A., Noroozi, O., Biemans, H. J. A., Mulder, M., & Banihashem, S. K. (2024). How does the type of online peer feedback influence feedback quality, argumentative essay writing quality, and domain-specific learning? *Interactive Learning Environments, 32*(9), 5459–5478. https://doi.org/10.1080/10494820.2023.2215822

van Popta, E., Kral, M., Camp, G., Martens, R. L., & Simons, P.R.-J. (2017). Exploring the value of peer feedback in online learning for the provider. *Educational Research Review, 20*, 24–34. https://doi.org/10.1016/j.edurev.2016.10.003

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor Package. *Journal of Statistical Software, 36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Viechtbauer, W., & Cheung, M.W.-L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods, 1*(2), 112–125. https://doi.org/10.1002/jrsm.11

Vogel, F., Wecker, C., Kollar, I., & Fischer, F. (2017). Socio-cognitive scaffolding with computer-supported collaboration scripts: A meta-analysis. *Educational Psychology Review, 29*(3), 477–511. https://doi.org/10.1007/s10648-016-9361-7

*Wichmann, A., Funk, A., & Rummel, N. (2018). Leveraging the potential of peer feedback in an academic writing activity through sense-making support. *European Journal of Psychology of Education*, *33*(1), 165–184. https://doi.org/10.1007/s10212-017-0348-7

Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting learners' agentic engagement with feedback: A systematic review and a taxonomy of recipience processes. *Educational Psychologist, 52*(1), 17–37. https://doi.org/10.1080/00461520.2016.1207538

Wu, Y., & Schunn, C. D. (2020). From feedback to revisions: Effects of feedback features and perceptions. *Contemporary Educational Psychology, 60*, 101826. https://doi.org/10.1016/j.cedpsych.2019.101826

Wu, Y., & Schunn, C. D. (2020). When peers agree, do students listen? The central role of feedback quality and feedback frequency in determining uptake of feedback. *Contemporary Educational Psychology, 62*, 101897. https://doi.org/10.1016/j.cedpsych.2020.101897

*Zeeb, H., Bürgermeister, A., Saalbach, H., Renkl, A. & Glogger-Frey, I. (2024). Effects of a digital support tool on student teachers' knowledge about, assessment of, and feedback on self-regulated learning. *Unterrichtswisssenschaft, 52*(1), 93–115. https://doi.org/10.1007/s42010-023-00184-z

*Zhao, Y., Zhang, F., Schunn, C. D., He, P., Di, L., & Zhao, Y. (2023). Feedback, feedback-on-feedback and re-feedback: Effects of written dialogic peer feedback on English as a foreign language writing. *Assessment & Evaluation in Higher Education, 49*(4), 523–538. https://doi.org/10.1080/02602938.2023.2278017

*Zheng, L. Q., Cui, P. P., Li, X., & Huang, R. H. (2017). Synchronous discussion between assessors and assessees in web-based peer assessment: Impact on writing performance, feedback quality, metacognitive awareness and self-efficacy. *Assessment & Evaluation in Higher Education*, *43*(3), 500–514. https://doi.org/10.1080/02602938.2017.1370533

Zheng, L., Zhang, X., & Cui, P. (2020). The role of technology-facilitated peer assessment and supporting strategies: A meta-analysis. *Assessment & Evaluation in Higher Education, 45*(3), 372–386. https://doi.org/10.1080/02602938.2019.1644603

*Zhu, W. (1995). Effects of training for peer response on students' comments and interaction. *Written Communication*, *12*(4), 492–528. https://doi.org/10.1177/0741088395012004004

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Julia Hornstein[1] · Melanie V. Keller[2] · Martin Greisel[1] · Markus Dresel[2] · Ingo Kollar[1]

✉ Julia Hornstein
julia.hornstein@uni-a.de

Melanie V. Keller
melanie.keller@uni-a.de

Martin Greisel
martin.greisel@uni-a.de

Markus Dresel
markus.dresel@uni-a.de

Ingo Kollar
ingo.kollar@uni-a.de

[1] Department of Educational Psychology, University of Augsburg, Universitätsstraße 10, 86159 Augsburg, Germany

[2] Department of Psychology, Universitätsstraße 10, 86159 Augsburg, Germany