Promoting future teachers' evidence-informed reasoning scripts: effects of different forms of instruction after problem-solving

Theresa Krause-Wichmann, Martin Greisel, Christina Wekerle, Ingo Kollar, Robin Stark

Check for updates

Promoting future teachers' evidence-informed reasoning scripts: Effects of different forms of instruction after problem-solving

Theresa Krause-Wichmann[1]*, Martin Greisel[2]*,
Christina Wekerle[2]*, Ingo Kollar[2] and Robin Stark[1]

[1]Department of Education, Saarland University, Saarbrücken, Germany, [2]Faculty of Philosophy and Social Sciences, Augsburg University, Augsburg, Germany

Pre-service teachers face difficulties when dealing with problem situations in the classroom if their evidence-informed reasoning script (EIRS) is not adequately developed. An EIRS might be promoted by demonstrating how to implement evidence-informed reasoning after a problem-solving activity on an authentic case. However, it is unclear what form of instruction is appropriate to promote pre-service teachers in the development of an EIRS. The present 2×3-factorial experimental intervention study investigated how different forms of instruction on functional procedures (example-free vs. example-based) and on dysfunctional procedures (without vs. example-free vs. example-based) affect the development of an EIRS. *N*=384 pre-service teachers worked on a written case vignette of a problem situation in a problem-solving phase, in which the crucial steps of the EIRS were prompted externally. In the subsequent instruction phase, the participants compared their own solution with an example-free or example-based instruction on functional procedures, which was either supplemented by an example-free or example-based instruction on typical dysfunctional procedures or not at all. The participants' learning success (declarative EIRS; near and far transfer problem-solving performance) and error awareness were assessed. The results revealed that the example-based instruction on functional procedures led to a higher learning success than the example-free instruction. Both forms of instruction on dysfunctional procedures improved learning success compared to learning without one. During learning, error awareness was higher for learners who worked with an example-free instruction on dysfunctional procedures. In order to promote the development of an EIRS in pre-service teachers, it is promising to provide instruction after problem-solving that presents a functional example of evidence-informed reasoning for the given problem and that also points out typical dysfunctional approaches to solving the problem. The results highlight the importance of selecting appropriate scaffolds in case-based learning approaches that aim to develop cognitive schemata. The mechanisms that explain when and why instructions on dysfunctional procedures work need to be further explored.

KEYWORDS

teacher education, evidence-informed reasoning, problem-solving prior to instruction, script theory, example-based learning, error-based learning

1. Introduction

1.1. Background and relevance

For teachers it is part of their daily teaching practice to be confronted with various problem situations in the classroom, for example, when students have difficulties in grasping the learning material or are unmotivated. Against the background of the constant demand for *evidence-informed teaching practice*, teachers are increasingly asked to base their decisions not (only) on subjective theories or experiential knowledge, but especially on educational theories and empirical findings (Joyce and Cartwright, 2020; Ferguson, 2021; Slavin et al., 2021). For example, when confronted with a student who is frustrated because of a poor grade, it might be appropriate to frame the feedback in terms of Weiner's (1986) attribution theory of motivation and emotion by attributing the performance to variable characteristics (e.g., the student's learning effort). Teachers must be able to understand theories and/or empirical findings, and to apply them in an appropriate and meaningful way. In particular, when reflecting on problem situations, the cognitive processes during problem-solving should be well selected and systematically carried out to avoid hasty and possibly dysfunctional decisions (Brown, 1987; Chen and Bradshaw, 2007; Jonassen, 2011; Csanadi et al., 2021). Therefore, this paper is concerned with how pre-service teachers can be supported in reflecting on problem situations through a systematic, coherent, and evidence-informed reasoning process.

1.2. Evidence-informed reasoning

Problem situations that teachers face in their daily practice can be distinguished into (a) problems that require immediate judgments and routines of action under time pressure, and (b) problems that allow for retrospective analysis and reflection; this latter type of problems provides an opportunity for evidence-informed reasoning (Renkl, 2022; Leuders et al., in press). In the educational context, evidence-informed reasoning can be defined as a process of thinking about a problem and forming an argument in a systematic and coherent way, underpinning the argument with educational theories and/or empirical findings (Csanadi et al., 2021; Wilkes and Stark, 2022). Evidence-informed reasoning tends not only to require knowledge of theories and empirical findings, but also knowledge of *what* actions or steps to take and *how* to take them to solve the problem at hand (van Gog et al., 2019). Research on teachers' reasoning skills (e.g., Seidel and Stürmer, 2014; Kiemer and Kollar, 2018; Kramer et al., 2021) suggests a sequence of four key activities that are useful in retrospectively reflecting on problem situations: Having identified a problem, the teacher must (1) reconstruct the problem by developing an understanding of which particular aspect of the given complex situation is actually the core of the problem (*problem description*). (2) The teacher must explain the problem by developing a causal model that represents relevant cause-and-effect relations (*problem explanation*). (3) The teacher must derive student-related consequences or target states by deducing from the previous step which alternative state should be aimed (*goal setting*). (4) The teacher must derive self-related consequences, i.e., concrete options for action that are suitable for

achieving the goals set in the previous step (*setting options for action*).

From a script-theoretical perspective (Schank and Abelson, 1977; Schank, 1999; Fischer et al., 2013), knowledge about *what* actions to perform and *how* to perform them is mentally organized in so-called *scripts*, which are a particular form of cognitive schemata. Teachers' knowledge about how to solve problem situations in the classroom can be conceptualized as a dynamic knowledge structure that guides teachers in solving a problem – an *evidence-informed reasoning script* (*EIRS*).

Several studies indicated that pre-service teachers and in-service teachers rarely display systematic, coherent, and evidence-informed reasoning: they show deficits in applying the crucial cognitive processes and/or evidence to the problem situation (e.g., Hetmanek et al., 2015; Lysenko et al., 2015; Yeh and Santagata, 2015; Kiemer and Kollar, 2018, 2021; Csanadi et al., 2021). These deficits could be explained not only by a lack of educational knowledge or affective barriers, such as negative beliefs regarding the usefulness of educational evidence (e.g., Dagenais et al., 2012; Lysenko et al., 2014; Kiemer and Kollar, 2021; Thomm et al., 2021), but also by an insufficiently developed EIRS (e.g., Kiemer and Kollar, 2018; Csanadi et al., 2021). Therefore, it is the responsibility of teacher education to find ways to support future teachers in developing an EIRS.

1.3. Fostering the development of an EIRS through authentic cases

When it comes to solving new, unfamiliar problem situations, teachers must be able to identify a problem, apply acquired knowledge and solve the problem systematically. To enable future teachers to tackle different problems based on a stable, well-developed EIRS, an important goal of teacher education is to teach for knowledge transfer. Previous research has extensively addressed the cognitive and situational aspects of learning and transfer, considering *what* is to be learned (e.g., abstract concepts, procedures), *to which situation or task* it is to be transferred (e.g., near transfer within a domain, far transfer beyond a domain), and *which instructional approach* is effective (e.g., problem-solving, metacognitive prompts; van Gog et al., 2019; Jacobson et al., 2020).

In terms of fostering transferable, evidence-informed reasoning skills in pre-service teachers, reflection on authentic and problematic case scenarios from pedagogical practice has proven valuable (e.g., Piwowar et al., 2018; Thiel et al., 2020; Helleve et al., 2021). From a cognitive perspective, case-based reasoning is beneficial for learning as it encourages learners to solve new problem situations by remembering previous situations and adapting their solutions (Kolodner, 1993). Despite the widespread use of case-based approaches, simply exposing pre-service teachers to complex problems from educational practice without further instructional guidance may not be sufficient to foster the development of an EIRS: Complex problems, such as classroom situations, place high demands on cognitive and metacognitive abilities, so that learners may perceive the difficulty of the problem situation as quite high; the capacity of working memory is likely to be overloaded, which may lead to cognitive overload (Ge and Land, 2003, 2004; van Merriënboer et al., 2003; Ge et al., 2005; Kirschner et al., 2006; van Merriënboer and Sweller, 2010; Jonassen, 2011). Therefore, learning with authentic

cases needs to be carefully instructed in teacher education to be beneficial for learning.

1.3.1. Supporting student learning

In terms of supporting student learning in problem-solving or reasoning tasks, instructional means that monitor cognitive processes and direct attention to critical aspects, such as external prompts, show promise (Ge and Land, 2003; Ge et al., 2005; Chen and Bradshaw, 2007; Wilkes et al., 2022). However, learners are not able to judge whether they have performed the requested activities in an appropriate way if they are only guided to solve the problem by external prompts (Spensberger et al., 2021). This means learners only know *which* actions they are supposed to perform in a templated manner, but they do not become aware of whether the way they performed the activities was *functional* (i.e., more likely to be correct) or *dysfunctional* (i.e., more likely to be incorrect). Therefore, it seems promising to help pre-service teachers not only to follow the sequence of the four reasoning steps mentioned above (i.e., *problem description, problem explanation, goal setting, setting options for action*) in problem-solving, but also to teach them *how* to perform these steps.

1.3.2. Instruction after problem-solving

One approach that could help to promote the acquisition and transfer of an EIRS is the *instruction after problem-solving approach* (*PS-I*; also referred to as *problem-solving prior to instruction*), which includes both an initial problem-solving phase and a subsequent instruction phase. In the initial problem-solving phase, learners attempt to solve a problem that requires the application of yet to-be-learned principles, concepts, or strategies, and often fail to solve the problem successfully; in the subsequent instruction phase, learners are explicitly taught the content to be learned (e.g., principles, concepts or strategies that should be applied to the given problem; e.g., Loibl et al., 2017; Sinha et al., 2020; Sinha and Kapur, 2021a).

Research on instruction after problem-solving in STEM fields. The benefits of PS-I have been empirically demonstrated and replicated in a variety of contexts, especially in comparison to approaches with direct instruction or instruction prior to problem-solving; PS-I has become particularly popular in STEM domains, with the goal of promoting conceptual learning and transfer (for an overview cf. Loibl et al., 2017; Sinha and Kapur, 2019, 2021b). For example, a typical problem-solving task used in PS-I for learning mathematical concepts is that students are given data of different athletes and asked to identify the most consistent athlete based on a mathematical calculation (e.g., Kapur, 2012, 2014). In the instruction phase, students are taught the canonical solution based on the mathematical concept.

Key mechanisms of PS-I. Three recent reviews have addressed the possible reasons of when and why PS-I is effective – or is not effective (Loibl et al., 2017; Sinha and Kapur, 2019, 2021b). These reviews indicate that the effectiveness of PS-I does not seem to be rooted in its individual components (i.e., problem-solving and instruction), but in the way they are combined and the sequencing of the phases. From these reviews, at least three key mechanisms can be derived as to why PS-I is conducive to learning, namely (1) that prior knowledge is activated and differentiated during problem-solving, (2) that learners' attention is directed to the principles, concepts or strategies to be learned in the instruction phase, and (3) that learners become aware of their dysfunctional procedures (i.e., errors) by questioning their own solutions (*error awareness;*

Loibl et al., 2017; Sinha and Kapur, 2019, 2021b). It should be noted that the implementation of both the problem-solving phase and the instruction phase differs across PS-I studies (Loibl et al., 2017; Sinha and Kapur, 2019, 2021b; Nachtigall et al., 2020). The popular *productive failure* approach, for which Sinha and Kapur (2021b) have formulated several fidelity criteria (e.g., providing problems that afford multiple representations, instruction building on students' solution), can be considered as a subtype of PS-I, but not all PS-I designs are examples of productive failure (for all fidelity criteria cf. Sinha and Kapur, 2021b). Overall, while the PS-I approach is very specifically characterized by its two phases, there is no single design of the two phases *per se* within the PS-I research; this is especially true for PS-I designs in non-STEM domains.

Instruction after problem-solving in fields beyond STEM. Given the large number of studies on PS-I in STEM fields, it is striking that the evidence on the impact of PS-I in less structured domains such as teacher education appears to be insufficient and inconsistent: Only a few studies investigated the effects of PS-I in less structured domains and these studies did not consistently indicate positive effects on learning (for an overview cf. Nachtigall et al., 2020). As different learning goals require different means of instruction, the inconsistent evidence could be explained by divergent learning goals (e.g., conceptual vs. procedural knowledge) and/or divergent design features (Loibl et al., 2017; Sinha and Kapur, 2019, 2021b; Nachtigall et al., 2020). For example, Schwartz and Bransford (1998) showed beneficial effects of PS-I with college students in the context of psychology, compared with instruction after reading or summarizing a text or with problem-solving without instruction. In the problem-solving phase of their PS-I condition, students had to analyze contrasting cases of data from simplified classical psychology experiments before engaging with a text or lecture on the relevant psychological phenomena. In a study by Glogger-Frey et al. (2015), who used PS-I in the domain of educational psychology with pre-service teachers to promote their abilities to assess learning strategies in learning journals, the PS-I condition was outperformed by a condition in which students studied a worked-out solution for the same problem-solving task before instruction. In the PS-I condition, participants were first given samples of learning journals written by (high-school) students and asked to develop criteria in order to assess the application of learning strategies. In the subsequent instruction phase, they were taught the evaluation criteria to be learned. However, the studies by Schwartz and Bransford (1998) and Glogger-Frey et al. (2015) differ from rather traditional PS-I studies (such as in STEM domains; Loibl et al., 2017; Sinha and Kapur, 2019, 2021b) regarding the control condition, which is that PS-I was not compared to instruction prior to problem-solving or direct instruction. Overall, it is difficult to draw conclusions about whether, when and why PS-I is also effective in domains that are rather less structured than STEM domains (such as teacher education).

Implications for the design of the problem-solving phase in less structured domains: Implementing structuring scaffolds. One design feature that is particularly different across PS-I studies (both in rather well- and less-structured domains) is the form of scaffolding in the problem-solving phase. In more traditional PS-studies, students received little or no support in solving a particular problem (e.g., Kapur, 2012, 2014). An advantage of unguided problem-solving is seen primarily in the fact that learners are given the opportunity to explore the problem, considering their own intuitive ideas (e.g., Sinha

et al., 2020). To specifically promote comprehensive exploration processes, Sinha et al. (2020) and Sinha and Kapur (2021a) offered so-called *failure-driven scaffolds* that explicitly encourage learners to explore the problem with suboptimal representations and solution paths. Other studies offered rather *success-driven scaffolds* that guide students, structure their problem-solving process, and thereby help them to perform better – at the expense of less opportunity to explore the problem. Examples of more success-driven scaffolds include contrasting cases (e.g., Schwartz and Bransford, 1998; Loibl and Rummel, 2014a; Glogger-Frey et al., 2015; Chase and Klahr, 2017), self-explanation prompts (e.g., Roll et al., 2012; Fyfe et al., 2014), interaction support (e.g., Roll et al., 2012; Westermann and Rummel, 2012) and accuracy feedback (e.g., Fyfe et al., 2014; for an overview see Loibl et al., 2017; Sinha and Kapur, 2019, 2021b; Nachtigall et al., 2020). The above-mentioned study by Sinha et al. (2020) indicated that failure-driven scaffolding is more effective in learning with the PS-I approach than success-driven scaffolding with high specificity (i.e., definite advice for the optimal solution), but similar to success-driven scaffolding with low specificity (i.e., external prompts or hints that structure the problem-solving process into subtasks, or tell students what to do, but not how to do it). Based on the extensive research on scaffolding, albeit mainly in STEM fields, it could be postulated that especially for complex problems in rather less structured domains, it might be promising to structure the problem-solving process using success-driven scaffolds in the form of low-specific problem-solving prompts (Jonassen, 1997, 2000, 2011; Chen and Bradshaw, 2007). In the learning environment of the present study, which aims to encourage pre-service teachers to apply a functional EIRS on problem situations in the classroom, learners were supported by the means of more success-driven prompts that structured their problem-solving process along the EIRS with low specificity, i.e., without suggesting precise procedures or instructing how to perform them.

Implications for the design of the instruction phase in less structured domains: The form of instruction deserves more attention. However, compared to the state of research on the implementation of guidance in the problem-solving phase of PS-I, only little research has addressed the design features of the instruction phase (Loibl et al., 2017; Sinha and Kapur, 2019, 2021b). In line with Sinha and Kapur (2021b), encouraging learners to compare their own solutions with a sample solution and its critical features in the instruction phase can be seen as a central part of the PS-I approach. Working through these critical features and becoming aware of specific knowledge gaps can encourage learners to rethink their mental models, trigger active processing of the content to be learned, and promote knowledge acquisition (Chi, 2000; Loibl and Rummel, 2014b; Sinha and Kapur, 2021b). In contrast to more well-structured STEM problems, there is rarely *one single* canonical, i.e., "correct" solution for problems teachers face in the classroom. There are usually several functional options for action and even more dysfunctional options. Against this background, we argue that the form of instruction deserves special attention when it comes to helping pre-service teachers to develop an EIRS. It seems important to consider not only the potential benefits of guidance in the problem-solving phase, but also to explore the question of what features of the instruction phase make learning with the PS-I approach beneficial, especially in rather less structured domains such as teacher education. We argue that pre-service teachers should not be taught (only) a single canonical solution for a problem

in the instruction phase, but especially how to apply a systematic, coherent, and evidence-informed problem-solving approach. This raises the question of how the instruction phase in PS-I can be designed to foster pre-service teachers' EIRS.

1.4. Designing instruction after problem-solving

1.4.1. Form of instruction on functional procedures

When there are several potentially "correct" options to solve a problem, as it is the case with most classroom problems, it is crucial to understand the rationale behind the problem and to apply powerful strategies or heuristics to tackle it (van Gog et al., 2004). If learners are presented with (only) an exemplary, worked-out solution to functionally deal with a given complex problem, there is a risk that learning will be hindered; learners might focus their attention on non-essential parts of the exemplary, worked-out solution (e.g., the specific wording) rather than on the underlying concepts, principles, or strategies (Renkl, 2002, 2017). It would be particularly precarious if learners misunderstand the meaning behind the exemplary, worked-out solution and do not become aware of their own dysfunctional procedures. In the worst case, not recognizing dysfunctional approaches could lead to learners internalizing and applying these approaches to problem situations in practice (Metcalfe, 2017).

Lange et al. (2021) therefore hypothesized that instructional explanations that do not include worked-out solutions (which they call *example-free instruction*) might be more effective for learning with complex problems than exemplary, functional worked-out solutions without instructional explanations. In their study, they examined the effects of such worked-out solutions without instructional explanations vs. example-free instruction on university students' critical thinking skills. While the worked-out solution only illustrated an exemplary, functional solution without instructional explanations, the example-free instruction provided extensive explanations for solving critical thinking problems. The example-free instruction has been proven to be superior to the worked-out solution without instructional explanations in promoting skill acquisition. Transferring Lange et al.' (2021) findings to teacher education, one could argue that it could be more appropriate to provide pre-service teachers with an example-free instruction after a problem-solving activity, explaining how to functionally manage the situation according to a normative EIRS in an abstract, general form.

On the other hand, research on example-based learning has revealed that examples are particularly promising to promote learning and transfer (van Gog and Rummel, 2010, 2018). Illustrating the content to be learned, such as abstract concepts, principles, or strategies by an example, can encourage learners to encode and interconnect both these abstract concepts, principles, or strategies and specific application possibilities; in this way, examples foster learners' ability to transfer acquired knowledge to new problem situations, and prevent them from acquiring so-called *inert knowledge* (i.e., knowledge that can be expressed but not applied to solve problems; Renkl et al., 1996; van Gog and Rummel, 2010, 2018; Renkl, 2017; Mayer, 2020). In example-based learning, examples are usually implemented in a way that concretely illustrates the application of the abstract concepts,

principles, or strategies that are to be explained to the learners; to this end, a written step-by-step sample solution is provided of how a particular problem can be solved in a functional way (commonly known as *worked example*; e.g., Atkinson et al., 2000; van Gog and Rummel, 2010, 2018). In terms of promoting the acquisition and transfer of the EIRS, pre-service teachers might benefit from an *example-based instruction after problem-solving* that combines both an abstract, general description of the normative EIRS with a worked example that illustrates a functional problem-solving approach according to the EIRS (not to be mistaken with the approach of using worked examples as preparatory activity *prior to instruction*; cf. Glogger-Frey et al., 2015). Providing students with both a description of the EIRS and an example of how to apply it would allow students to better understand the rationale of the EIRS; moreover, they may become more easily aware of their own dysfunctional procedures (Loibl et al., 2017).

Thus, the question arises whether instruction after problem-solving should describe the functional operations of the above four reasoning steps in a still abstract, general form (i.e., *example-free instruction on functional procedures*), or in a worked-out form, in which it is – in addition to a general description of the normative EIRS – concretely illustrated how to solve the given problem by applying the operations of the EIRS (i.e., *example-based instruction on functional procedures*).

1.4.2. Form of instruction on dysfunctional procedures

Providing instruction on functional procedures helps to build knowledge about *what to do best* when faced with a particular problem or task (Oser et al., 2012; Renkl, 2017). Learners could benefit not only from instruction that focuses on *best practice*, but also from instruction that focuses on *dysfunctional practice* (e.g., Loibl and Rummel, 2014b; Loibl and Leuders, 2018, 2019). In terms of reflecting on problem situations in the classroom, a typical example of dysfunctional practice is that the third step of the EIRS *goal setting* is skipped. In other words, when a teacher has analyzed the problem (i.e., EIRS step 1: *problem description*) and its possible reasons or consequences (i.e., EIRS step 2: *problem explanation*) in a functional way, the teacher already formulates concrete options for action (i.e., EIRS step 4: *setting options for action*). It would have been important to first consider target states for the student to be achieved from the perspective of theory or empirical findings (i.e., EIRS step 3: *goal setting*) in order to avoid jumping to conclusions. Another typical dysfunctional practice is that the EIRS step *goal setting* is implemented in an inappropriate way: For example, even if a student-related target state is formulated before concrete options for action are determined (i.e., EIRS step 4), the target state is not coherent with the previous explanation of the problem and, therefore, may not contribute to solving the problem.

Tracing how a dysfunctional procedure differs from a functional one and understanding why a procedure is dysfunctional could help learners to correctly update schemata of functional procedures and to create schemata of dysfunctional procedures (i.e., *negative knowledge*; Oser et al., 2012). When learners thoroughly elaborate the features of dysfunctional procedures, they are more likely to address their own knowledge gaps (e.g., Große and Renkl, 2007; Durkin and Rittle-Johnson, 2012; Barbieri and Booth, 2020).

Error-based learning is criticized because it risks learners internalizing dysfunctional procedures (Metcalfe, 2017). It is important that learners are enabled to reflect on dysfunctional procedures in comparison to corresponding functional procedures (Durkin and Rittle-Johnson, 2012; Oser et al., 2012). In several PS-I studies demonstrating the benefits of presenting dysfunctional procedures learners were guided to compare both functional and dysfunctional solution attempts in whole-class discussions, i.e., only the teacher/instructor built upon learners' typical dysfunctional procedures and contrasted their features with those of functional procedures (e.g., Kapur, 2012, 2014; Kapur and Bielaczyc, 2012; Loibl and Rummel, 2014b). Loibl and Leuders (2018, 2019), therefore, investigated learners' individual cognitive processes during instruction. Both studies underline the advantages of providing both a solution of a functional procedure and typical dysfunctional solution attempts, supplemented by comparison prompts, as opposed to providing a solution of a functional procedure and typical dysfunctional solution attempts without comparisons prompts or only solutions of a functional procedure. Moreover, the study by Loibl and Leuders (2019) indicated that the elaboration of the dysfunctional procedures seemed to mediate this effect.

Looking at the PS-I approach from a conceptual change perspective (e.g., Vosniadou, 2013, 2019), one could argue that it might be promising for students' error awareness not to prompt the comparison of the functional procedure and typical dysfunctional procedures with each other, but of *their own* procedure with both a functional procedure and typical dysfunctional procedures. By comparing their own solution approach with both a solution of a functional procedure and of typical dysfunctional procedures, students might become more aware of the appropriateness of their own approach. When learners are explicitly encouraged to recognize dysfunctional procedures in their own solution, they can reorganize their mental schemata (Posner et al., 1982), and, as a result, be prevented from internalizing dysfunctional solution approaches (Metcalfe, 2017). A study by Heemsoth and Heinze (2016), for example, indicated the benefits of prompted reflection on the rationale behind one's own dysfunctional procedures (i.e., error-centered reflection) over the reflection on the correct solution corresponding to one's own dysfunctional procedures (i.e., solution-centered reflection) after problem-solving on knowledge acquisition. The authors explained the effect by suggesting that error-centered reflection would lead to more elaborated learning. Through error-centered reflection, students might have become more aware of the extent to which the procedures they used fostered or hindered the problem-solving process (Heemsoth and Heinze, 2016).

It remains unclear what form of instruction on dysfunctional procedures as a complement to instruction on functional procedures is appropriate to promote error awareness and the development of an EIRS in pre-service teachers. On the one hand, specific exemplification of *how not to do something* using *dysfunctional* (or *erroneous*) *examples* might encourage learners to identify, comprehend, explain, and/or remedy own dysfunctional procedures by referring to underlying concepts, principles, or strategies (e.g., Große and Renkl, 2007; Durkin and Rittle-Johnson, 2012; Barbieri and Booth, 2020). On the other hand, students might not benefit from exemplifications, if their own solution approach does not resemble the dysfunctional procedures presented; specific exemplifications of other students' dysfunctional procedures that have not been used by the learners themselves might even distract them from becoming aware of the correctness of their own approach (Loibl and Leuders, 2019).

Overall, it is unclear whether pre-service teachers learning with instruction on functional procedures after problem-solving would benefit from supplementary instruction on typical dysfunctional procedures of other pre-service teachers – be it in the form of an abstract, general description of typical dysfunctional procedures (i.e., *example-free instruction on dysfunctional procedures*) or in a worked-out form that – in addition to such an abstract, general description – also presents a specific exemplification of the described procedures for the given problem (i.e., *example-based instruction on dysfunctional procedures*).

1.5. Research questions and hypotheses

In the present experimental intervention study, pre-service teachers learned how to analyze educational problems in a systematic, coherent, and evidence-informed way, based on the PS-I approach. The study focused on the effects of different forms of instruction: On the one hand, we investigated which form of instruction on functional procedures (i.e., *example-free instruction on functional procedures* vs. *example-based instruction on functional procedures*) would be more suitable for pre-service teachers' ability to deal with problem situations according to the EIRS. Secondly, we aimed to find out whether and in what form instruction on functional procedures should be complemented by instruction on typical dysfunctional procedures (i.e., *without any instruction on dysfunctional procedures* vs. *example-free instruction on dysfunctional procedures* vs. *example-based instruction on dysfunctional procedures*). We formulated the following research questions and hypotheses:

Research Question 1. To what extent do different forms of instruction on functional procedures (i.e., *example-free* vs. *example-based*) and on dysfunctional procedures (i.e., *without* vs. *example-free* vs. *example-based*) affect pre-service teachers' learning success, i.e., the development of an EIRS (i.e., *the declarative EIRS*) and its application in similar and unfamiliar problem situations (i.e., *near and far transfer problem-solving performance*)?

> *Hypothesis 1*: Against the background of the benefits of examples (e.g., van Gog and Rummel, 2010, 2018), we expected that the example-based instruction on functional procedures would lead to a higher learning success than learning with the example-free instruction on functional procedures. We further expected that learning with the example-based instruction on dysfunctional procedures would be superior to learning with the example-free instruction and without any instruction on dysfunctional procedures. Yet, the example-free instruction on dysfunctional procedures should still work better than no instruction on dysfunctional procedures. We further hypothesized that a combination of the example-based instruction on functional procedures and dysfunctional procedures would lead to the highest learning success.

Research Question 2: To what extent do different forms of instruction on functional procedures (i.e., *example-free* vs. *example-based*) and on dysfunctional procedures (i.e., *without* vs. *example-free* vs. *example-based*) affect pre-service teachers' *error awareness*, with special emphasis on the written comparison of the students' solution and the instruction during learning?

> *Hypothesis 2*: We assumed that learning with the example-based instruction on functional procedures would have a greater potential to help learners to become aware of the correctness of their own approach than learning with the example-free instruction on functional procedures. We postulated that the example-based instruction on dysfunctional procedures would promote learners' error awareness more than the example-free instruction or no instruction on dysfunctional procedures. Learning with the example-free instruction on dysfunctional procedures should still be superior to learning without any instruction on dysfunctional procedures to promote students' error awareness. The combination of example-based instruction on functional and dysfunctional procedures was expected to be superior to all the other conditions in terms of promoting error awareness.

Research Question 3. To what extent is the postulated effect of the form of instruction on functional procedures on near and far transfer problem-solving performance serially mediated by pre-service teachers' error awareness and the declarative EIRS, moderated by the different forms of instruction on dysfunctional procedures (i.e., *without*, *example-free*, and *example-based*)?

> *Hypothesis 3*: Following Loibl and Leuders (2019), we expected that the hypothesized effect of the form of instruction on functional procedures on near and far transfer problem-solving performance would be serially mediated by the participants' error awareness and the declarative EIRS for the three different forms of instruction on dysfunctional procedures. The instruction on dysfunctional procedures is viewed as complementary to the instruction on functional procedures, and thus is conceptualized as a moderator. We assumed that (a) the form of instruction on functional procedures would influence the participants' error awareness. Subsequently (b), the higher the participants' error awareness, the more pronounced their declarative EIRS should be. Finally (c), the better the participants' declarative EIRS would be developed, the better the students should be able to solve near and far transfer problems. Further, we postulated that (d) these associations would be moderated by the form of instruction on dysfunctional procedures, respectively (for the postulated moderated serial mediation, see Figure 1).

2. Method

2.1. Participants and design

$N = 384$ pre-service teachers ($M_{Age} = 27.72$, $SD = 3.54$; 76% female) participated as part of their regular university courses. On average, students were in their third semester ($M_{Sem} = 3.43$, $SD = 1.64$). It was mandatory to take part in the training elements of the study, but it was voluntary to participate in the data collection. As no one opted out, the full sample was included in the analyses.

In a randomized 2×3-factorial between-subjects design, the factors *form of instruction on functional procedures* and *form of instruction on dysfunctional procedures* were varied, resulting in six
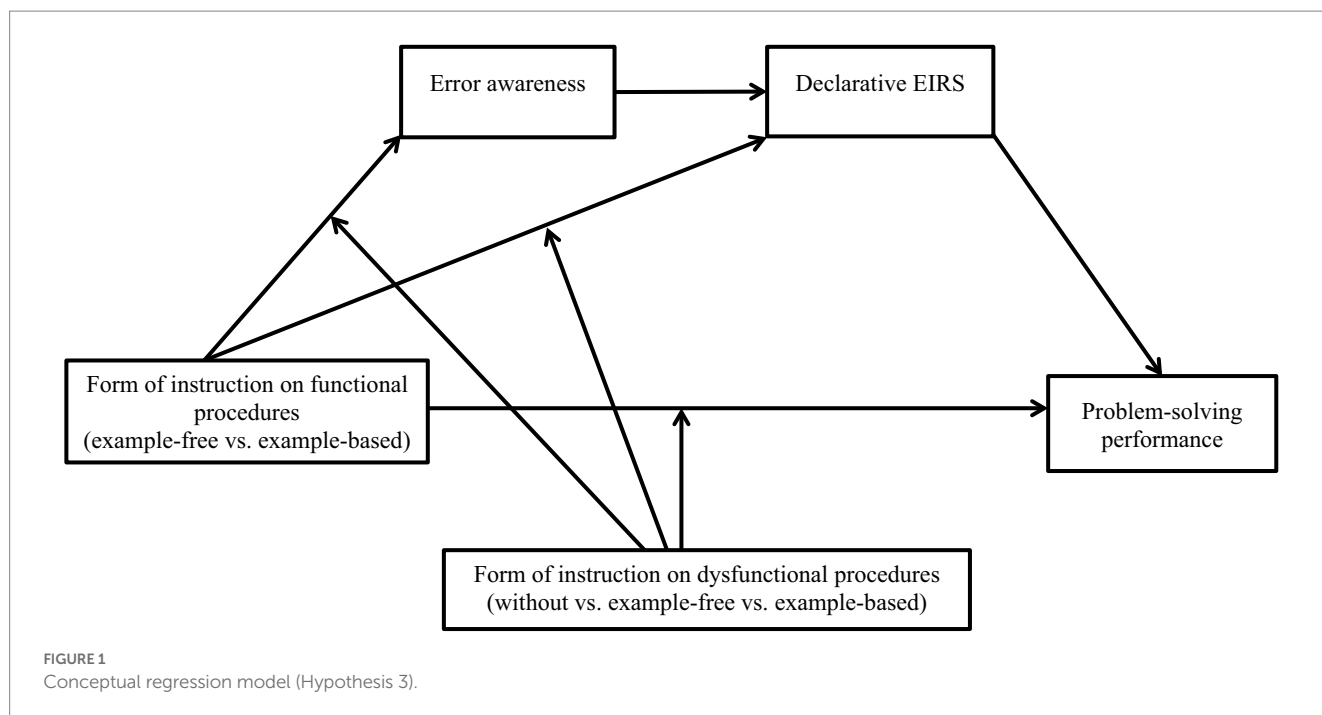
FIGURE 1
Conceptual regression model (Hypothesis 3).

TABLE 1 2×3-factorial, experimental design.

Factor 1: Form of instruction on functional procedures	Factor 2: Form of instruction on dysfunctional procedures		
	Without [0]	Example-free [D⁻]	Example-based [D⁺]
Example-free [F⁻]	F⁻ 0 ($n = 64$)	F⁻ D⁻ ($n = 64$)	F⁻ D⁺ ($n = 64$)
Example-based [F⁺]	F⁺ 0 ($n = 64$)	F⁺ D⁻ ($n = 64$)	F⁺ D⁺ ($n = 64$)

experimental conditions with $n = 64$ participants each (Table 1). With respect to the *form of instruction on functional procedures*, we varied whether students received either an *example-free* [F⁻] or an *example-based* [F⁺] instruction. Regarding the *form of instruction on dysfunctional procedures,* the participants were either not provided with any instruction on dysfunctional procedures [0], with an example-free [D⁻], or with an example-based instruction [D⁺]. A power analysis for 2×3-ANOVA with $\alpha = 0.05$, $1 - \beta = 0.80$ indicated that the sample size of $N = 384$ would be sufficient to identify small- to medium-sized effects.

2.2. Procedure and material

The study comprised two online sessions that were conducted independently and individually on a private computer with no time limit. An overview of the entire procedure is presented in Figure 2. In the first session, 2 weeks before the start of the training, participants answered a pre-questionnaire on socio-demographic data and a pre-test. The second session consisted of two parts: (1) the training with the actual intervention (i.e., problem-solving phase and instruction phase) and (2) a post-test.

The training followed the PS-I approach: in the initial problem-solving phase, participants worked through a written problematic classroom case scenario of 200 words using a written summary of

corresponding educational theories and empirical findings. The case scenario involves a student who is angry at her mathematics teacher because, in her opinion, a test she failed had been too difficult. As a result, she declares that she does not want to study for the subject anymore. The summary of educational theories and findings focused on the attribution theory of motivation and emotion (Weiner, 1986), control-value theory of achievement emotions (Pekrun, 2006), achievement goal theory (Wigfield et al., 2016), and corresponding empirical findings. To reduce the complexity of the case analysis, all participants received success-driven prompts that structured the problem-solving process along the EIRS with low specificity (cf. *Instruction after Problem-Solving*): Per prompt, learners were instructed to perform one of the four EIRS steps (i.e., *problem description, problem explanation, goal setting, setting options for action*). In doing so, the participants were not instructed *how* to perform the steps, i.e., no precise procedures or operations were suggested (e.g., for *goal setting*: "Please set general goals to improve the situation"). In the subsequent instruction phase, participants were asked to compare their own solution with the provided instruction, to explain what they did the same or similar, and what they did differently or incorrectly. The instructions explained how to perform the operations of the EIRS in different ways that varied depending on condition (cf. *Participants and Design* and *Operationalization of the Independent Variables*). After the training phase, the post-test was administered.

FIGURE 2
Procedure.
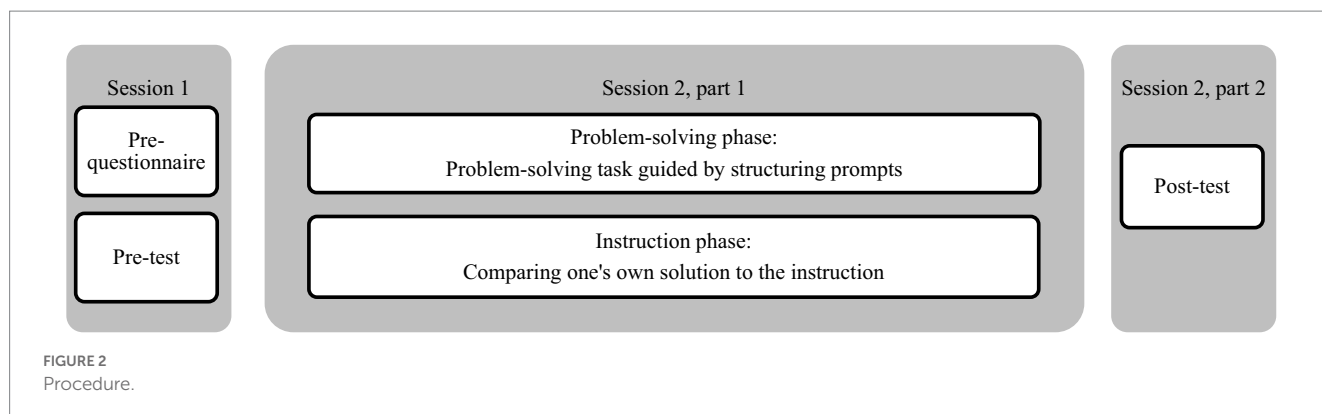
2.3. Operationalization of the independent variables

After the problem-solving phase, the participants with the *example-free instruction on functional procedures* received an instruction that described the crucial operations of a normative EIRS (i.e., how to ideally solve problem situations in a functional evidence-informed manner) in an abstract, general form, without any detailed elaboration of the EIRS for the given problem. For example, regarding the third step of the EIRS *goal setting*, the instruction focused on the two crucial operations of deriving student-related target states, and of connecting these relations with evidence (cf. Table 2).

The participants with the *example-based instruction on functional procedures* learned with an instruction that comprised the above-mentioned abstract description of the crucial operations of a functional EIRS, and a detailed "good practice" elaboration of these operations for the previous problem-solving task. For the step *goal setting*, for instance, it was concretely worked out which student-related goals could be derived coherently from the previously determined cause-and-effect relations, taking into account the evidence presented in the summary (cf. Table 2).

In the conditions that combined the instruction on functional procedures with an example-free instruction on dysfunctional procedures, the participants were also presented with several typical dysfunctional problem-solving procedures of pre-service teachers in an abstract form. Yet, this description did not include a precise elaboration of these typical dysfunctional procedures that might be applied to the problem in the previous training task. Regarding the step *goal setting*, for instance, students were told that pre-service teachers often select target states that do not fit to the cause-and-effect relations established before (cf. Table 2).

In the conditions that combined the instruction on functional procedures with an example-based instruction on dysfunctional procedures, participants also received the above-mentioned abstract, general description of typical dysfunctional procedures, along with a detailed elaboration of these dysfunctional procedures for the problem situation of the previous training task. In the step *goal setting*, for example, a typical student solution was illustrated that did not address the actual cause-effect-relations of the problem (cf. Table 2).

In the two conditions without any instruction on dysfunctional procedures, participants received only the assigned instruction on

functional procedures (either in an example-free or an example-based version, depending on condition).

2.4. Measures

2.4.1. Dependent measures

In the post-test, three indicators of knowledge acquisition were measured: (a) declarative EIRS, (b) near-transfer problem-solving performance, and (c) far-transfer problem-solving performance.

Declarative EIRS. To measure the participants' declarative EIRS, they were asked to indicate how they would proceed when analyzing problematic classroom situations by describing in an open-ended format how they would implement the individual steps in detail. Two coders, who were blind to condition, were trained to code the participants' answers. The answers were coded with respect to whether they included the operations that define the EIRS (i.e., *problem description, problem explanation, goal setting, setting options for action*). To calculate inter-rater reliability, 10 % of the sample were double-coded by the two independent coders, with satisfactory inter-rater reliability (*Cohen's* $\kappa > 0.80$). Students were able to reach a maximum of five points for the definition of the four EIRS steps (two of the four EIRS steps, i.e., *problem explanation* and *goal setting*, also include an appropriate reference to evidence, for which 0,5 points each were awarded; cf. Table 3).

Near transfer problem-solving performance. Near transfer problem-solving performance was measured by an analysis of the participants' written analyses of a case vignette that described a problem similar to the training scenario in 200 words. Essentially, the problem was that a student is dissatisfied with the grade in the test and while rethinking the main reasons for the grade (i.e., lack of talent), he finally concludes that future learning would be a waste of time. The students were asked to analyze the situation along the EIRS with the aid of educational evidence, while the summary of corresponding educational theories and empirical findings was not provided anymore. Since there is more than one possible "correct" option to handle a complex problem situation, the answers were coded with respect to whether the participants had appropriately applied the operations that define the EIRS (i.e., *problem description, problem explanation, goal setting, setting options for action*). In other words, two coders assessed *how* the students reflected on the problem and whether it was done in a systematic, coherent, and evidence-informed manner, based on the operations of the EIRS taught in the training, or not. The right column of Table 3 presents a coding

TABLE 2 Exemplary illustration of the different forms of instruction for the step *goal setting* (translated and abbreviated).

Condition	Instruction on functional procedures		Instruction on dysfunctional procedures		
	Example-free	Example-based	Without	Example-free	Example-based
	Concluding from the previous analysis and based on the evidence (theories and findings) you received, you need to derive student-related consequences, i.e., target states. In other words, you need to consider what alternative state should be targeted for the situation by starting at the adjusting screws of the previously identified negative cause-effect relations, so that positive cause-effect relations are established. Thereby, theoretical terms/evidence must be assigned appropriately.	[Text of respective example-free instruction, cf. left column.] In line with Weiner's (1986) attribution theory and Pekrun's (2006) control-value theory, Mr. Schuster must ensure that Sarah develops an internally variable attribution pattern and a higher control appraisal. Sarah further needs to realize that she can improve through focused learning (learning goal orientation; Wigfield et al., 2016). In order to do so, however, she needs to feel a sense of control so that she does not become frustrated. Moreover, Sarah should view poor grades less as a failure and more as an opportunity to close the identified gaps.		Typical errors of other students: (1) Immediately after the analysis, concrete options for teacher action are already formulated without deriving student-related target states that should be achieved from the perspective of the theory or empirical findings. (2) The derived target state does not fit the previously explained cause-effect relation.	[Text of respective example-free instruction, cf. left column.] (1) The teacher might encourage Sarah that she can get a better grade next time and give her individual exercises to do at home. (2) In order for Sarah to stop being frustrated, her grades need to improve.
F⁻ 0	x		x		
F⁻ D⁻	x			x	
F⁻ D⁺	x				x
F⁺ 0		x	x		
F⁺ D⁻		x		x	
F⁺ D⁺		x			x

example for the near transfer problem-solving performance. The coders, who were again blind to condition, reached satisfactory inter-rater reliability in 10 % of the sample (*Cohen's* κ > 0.80). Participants were able to score up to five points for correctly applying each operation of the EIRS to the problem situation.

Far transfer problem-solving performance. A second scenario measuring far transfer presented a completely novel situation, which had to be analyzed along the EIRS with the help of a summary of corresponding theories and empirical findings that was different from the training (self-determination theory; Ryan and Deci, 2000). In the 120-words-scenario, a science teacher fails to motivate his students to actively participate in his lesson. Two coders (blind to condition) estimated with satisfactory inter-rater reliability (in 10 % of the sample; *Cohen's* κ > 0.80) whether the four reasoning steps were correctly implemented or not. As with the near transfer problem-solving performance, participants could receive a maximum of five points (cf. Table 3).

To measure the participants' *error awareness* as a process variable in the instruction phase, we asked the participants to compare their own solution to the provided instruction, and to state what they did differently or what similarly when they were working on the problem. Two independent coders rated whether the participants had assessed

their approach correctly, i.e., whether the students identified their functional (i.e., rather correct) and dysfunctional (i.e., rather incorrect) solution steps in terms of the EIRS (max. 5 points; cf. Table 3). Again, 10 % of the sample were double-coded. Inter-rater reliability was satisfying (*Cohen's* κ > 0.80).

2.4.2. Control measures

Prior problem-solving performance. To capture the participants' prior problem-solving performance, the pre-test presented a 200-words-scenario, which resembled the scenario measuring near transfer problem-solving performance. Participants were asked to analyze the scenario using educational theories and/or empirical findings, if possible, but they were not given a summary of corresponding evidence at this point. As with the post-test, learners could score up to max. 5 points (cf. Table 3). Ten percent of the sample were double-coded with satisfactory inter-rater reliability (*Cohen's* κ > 0.80). Reliability was sufficient (*Cronbach's* α = 0.70).

Beliefs regarding the usefulness of educational evidence. In addition, participants had to rate five items measuring their beliefs regarding the usefulness of educational evidence (adapted from Wagner et al., 2016; e.g., "Educational knowledge is helpful to make good teaching decisions"; 1 = not at all true, 5 = very much true; Cronbach's α = 0.78).

TABLE 3 Coding scheme and coding example (translated).

EIRS operations (declarative EIRS)	Max. Points	Coding example for the application of the EIRS operations (near transfer problem-solving performance)	Coding (Points)
(1) Problem description: The core of the problem must be described.	1	The core of the problem is the lack of motivation to deal with the learning content. The student perceives help as annoying and rejects it.	1
(2) Problem explanation: a) A correct model of cause-and effect relations of the problem situation must be formulated…	1	According to Weiner and Pekrun, the student attributes his failures internally and stably, i.e., he attributes his failures to a lack of talent. Consequently, he does not see any sense in trying harder and dealing intensively with the contents, because he sees himself as not gifted. In the course of this, resignation sets in and the help of others is rejected. Here possibly a work avoidance goal orientation (goal orientation theory) threatens, in order to escape the shame that the student feels. The student does not see the everyday relevance of the topic, which affects additionally negatively his motivation.	1
b) …and theoretical terms/evidence must be assigned appropriately.	0,5		0,5
(3) Goal setting: a) Student-related target states must be derived immediately logical from the result in cause-and-effect relations of the problem…	1	In line with the provided evidence on the benefits of reattribution, reattribution should take place on the part of the student in order to improve his self-concept. He should learn that through stronger effort, he will achieve better grades (attribution: internal and variable). The goal orientation should also change, towards a learning goal orientation.	1
b) …and theoretical terms/evidence must be assigned appropriately.	0,5		0,5
(4) Setting options for action: Concrete options for action must be directly derived from the target states.	1	The teacher could conduct partner dictations, which are corrected by the partner. In this way, the students give a positive value to the task and have a high level of control, since they can control each other's tasks. Consequently, they might feel pleasure in solving the task.	0,5

2.5. Analytic strategy

As level of significance, α = 0.05 was applied for all global tests of significance. To answer Research Question 1, 2×3-factorial multivariate analyses of variance (MANOVAs) with the experimental factors *form of instruction on functional procedures* and *form of instruction on dysfunctional procedures* were calculated for the three interrelated learning success measures. To check effects on *error awareness* (Research Question 2), 2×3-factorial analyses of variance (ANOVAs) were computed. Planned contrast tests were performed (Research Question 1 and 2) to examine the presumed differences between the single conditions for significant main effects of the factor *form of instruction on dysfunctional procedures* and for significant interaction effects. If the ANOVA had revealed a significant main effect of the *form of instruction on dysfunctional procedures*, six contrasts tests were calculated that analyzed the postulated between-group differences in dependence of the three individual factor levels (i.e., without [0], example-free [D⁻] example-based [D⁺]).[1] In case of a significant interaction effect, the postulated superiority of the combination of the example-based instruction on functional procedures and the example-based instruction on dysfunctional procedures was checked by comparing this condition with all other conditions.[2] To account for potential alpha error inflation, the Bonferroni-Holm correction was applied.

Hayes' (2021) PROCESS macro for SPSS version 4.0 was used to examine the postulated mediation effect of the form of instruction on functional procedures on the near and far transfer problem-solving performance, moderated by the different forms of instruction on dysfunctional procedures (Research Question 3; Figure 1). Two moderated serial mediation models were computed with two mediators in a row (mediator 1: error awareness, mediator 2: declarative EIRS; Figure 1). The form of instruction on dysfunctional procedures was conceptualized as moderator. As dependent measure, the near transfer problem-solving performance was used in Model 1, and the far transfer problem-solving performance in Model 2. For mediation analyses, all mediating and dependent variables were z-standardized. To test the significance of indirect effects, we used 10,000 bootstrapped samples. We interpreted the indirect effect size as significant if its 95%-confidence interval did not include zero.

3. Results

3.1. Preliminary analyses

There were no *a-priori* differences between the experimental conditions regarding the control measures prior problem-solving

---

1   In case of a significant main effect of the *form of instruction on dysfunctional procedures*, the following conditions were compared with each other: (1) *F⁺ D⁺* vs. *F⁺ D⁻*, (2) *F⁺ D⁺* vs. *F⁺ 0*, (3) *F⁺ D⁻* vs. *F⁺ 0*, (4) *F⁻ D⁺* vs. *F⁻ D⁻*, (5) *F⁻ D⁺* vs. *F⁻ 0*, (6) *F⁻ D⁻* vs. *F⁻ 0*).

2   In case of a significant interaction effect, the condition *F⁺ D⁺* was compared to (1) *F⁺ D⁻*, (2) *F⁺0*, (3) *F⁻ D⁻*, (4) *F⁻ D⁺*, (5) *F⁻ 0*.

performance, $F(5, 378) = 0.99$, $p = 0.424$, and beliefs regarding the usefulness of educational evidence, $F(5, 378) = 0.737$, $p = 0.595$ (for means and standard deviations cf. Table 4).

3.2. Learning success (Research Question 1)

The means and standard deviations of the three learning success measures (i.e., declarative EIRS, near transfer problem-solving performance, far transfer problem-solving performance) for the six experimental conditions are displayed in Table 4. Overall, it is remarkable that all values were rather far away from the theoretical maximum to be reached in the post-test. The correlations between the three learning success measures (i.e., declarative EIRS, near transfer problem-solving performance, far transfer problem-solving performance) are presented in Table 5. The correlations were small to moderate and positive (Table 5). The results of the MANOVAs, the subsequent ANOVAs and the planned contrasts are reported below. The results of the planned contrast tests for a significant main effect of the form of instruction on dysfunctional procedures are presented in Table 6, those for a significant interaction effect in Table 7.

The 2×3-factorial MANOVA, using the *form of instruction on functional procedures* as well as the *form of instruction on dysfunctional procedures* as between-subject factors, and the three learning success measures as dependent measures revealed no significant interaction effect, $\Lambda = 0.98$, $F(6, 752) = 1.60$, $p = 0.144$. There was a significant main effect with moderate effect size for the form of instruction on functional procedures, $\Lambda = 0.93$, $F(3, 376) = 9.22$, $p < 0.001$, $\eta_p^2 = 0.07$. There was also a significant main effect with small effect size for the form of instruction on dysfunctional procedures, $\Lambda = 0.94$, $F(6, 752) = 3.69$, $p < 0.001$, $\eta_p^2 = 0.03$.

Concerning the *declarative EIRS*, in line with the MANOVA, the subsequent 2×3-factorial ANOVA showed no interaction effect, $F(2, 378) = 1.50$, $p = 0.225$. The results showed a small significant main effect of the form of instruction on functional procedures, $F(1, 378) = 5.25$, $p = 0.023$, $\eta_p^2 = 0.014$, indicating that participants working with the example-based instruction reproduced more operations of the EIRS than those participants working with the example-free instruction (Table 4). Further, there was a small significant main effect of the form of instruction on dysfunctional procedures, $F(2, 378) = 8.09$, $p < 0.001$, $\eta_p^2 = 0.041$. The planned contrasts showed that participants who were provided with any kind of instruction on dysfunctional procedures (i.e., example-free or example-based) outperformed those who were not, but these differences were only significant in the conditions with the example-based instruction on functional procedures (Tables 4, 6: Contrasts 2 and 3). For participants who learned with the example-free instruction on functional procedures the superiority of the two forms of instruction on dysfunctional procedures could only be observed at a descriptive level (Tables 4, 6: Contrasts 5 and 6). There were no significant differences between the conditions with the example-based instruction on dysfunctional procedures and the respective conditions with the example-free instruction on dysfunctional procedures with regard to influencing the acquisition of the EIRS (Table 6: Contrasts 1 and 4).

A similar picture emerged regarding the *near transfer problem-solving performance*. Consistent with the MANOVA, the interaction between both factors failed to reach significance, $F(2, 378) = 0.94$, $p = 0.392$. There was a small significant main effect of the form of

instruction on functional procedures, $F(1, 378) = 7.04$, $p = 0.008$, $\eta_p^2 = 0.018$, indicating that the example-based instruction was superior to the example-free instruction in terms of fostering learning success (Table 4). Results further revealed a small significant main effect of the form of instruction on dysfunctional procedures, $F(2, 378) = 6.51$, $p = 0.002$, $\eta_p^2 = 0.033$. Contrast tests showed that students who worked with the example-free instruction on dysfunctional procedures outperformed those students who worked without any instruction on dysfunctional procedures, but only in the condition with the example-based instruction on functional procedures (Tables 4, 6: Contrast 3). Although the other calculated contrasts were not significant, it should be noted that for the participants in the condition with the example-free instruction on functional procedures, the latter finding could be observed at a descriptive level (Tables 4, 6: Contrast 6). The superiority of learning with the example-based instruction on dysfunctional procedures over learning without any instruction on dysfunctional procedures could be revealed only on a descriptive level, for both the conditions with the example-free and example-based instruction on functional procedures (Tables 4, 6: Contrasts 2 and 5).

Although the interaction effect of the MANOVA was not significant, the ANOVA for the *far transfer problem-solving performance* revealed a small significant interaction effect; it indicated that the example-based instruction on functional procedures was only superior, if it was combined with one of the instructions on dysfunctional procedures, $F(2, 378) = 3.27$, $p = 0.039$, $\eta_p^2 = 0.017$ (Figure 3; Table 4). The planned contrast tests revealed that participants who learned with the example-based instruction on functional procedures in combination with the example-based instruction on dysfunctional procedures showed a better performance than participants in all the other conditions (Tables 4, 7: Contrasts 2–5) – except for the condition that combined the example-based instruction on functional procedures with an example-free instruction on dysfunctional procedures; here, the result was on a descriptive level (Table 7: Contrast 1). Since the interaction was ordinal, main effects can be interpreted (Figure 3). The main effect of the form of instruction on functional procedures was significant with moderate effect size, $F(1, 378) = 24.54$, $p < 0.001$, $\eta_p^2 = 0.061$. Participants who received the example-based instruction showed a higher learning success than those who received the example-free instruction (Table 4). However, there was no significant main effect of the form of instruction on dysfunctional procedures, $F(2, 378) = 1.06$, $p = 0.347$.

3.3. Error awareness (Research Question 2)

Regarding *error awareness*, neither the interaction effect, $F(2, 378) = 1.68$, $p = 0.188$, nor the main effect of the instruction on functional procedures, $F(1, 378) = 2.07$, $p = 0.151$ reached the level of statistical significance in the 2×3 ANOVA. Yet, there was a small significant main effect of the form of instruction on dysfunctional procedures, $F(2, 378) = 7.36$, $p < 0.001$, $\eta_p^2 = 0.038$. The contrast tests revealed that the participants who learned with the example-free instruction on dysfunctional procedures as a supplement for the example-based instruction on functional procedures had a higher error awareness than those who received the example-based instruction on dysfunctional procedures as a supplement (Tables 4, 6: Contrast 1). All other contrasts (Table 6) did not reach statistical significance. However, it should be noted, that on a descriptive level,

TABLE 4 Means and standard deviations for all variables depending on the experimental condition.

Variable	Condition					
	Example-based instruction on functional procedures			Example-free instruction on functional procedures		
	Example-based instruction on dysfunctional procedures	Example-free instruction on dysfunctional procedures	Without instruction on dysfunctional procedures	Example-based instruction on dysfunctional procedures	Example-free instruction on dysfunctional procedures	Without instruction on dysfunctional procedures
	(F⁺ D⁺)	(F⁺ D⁻)	(F⁺ 0)	(F⁻ D⁺)	(F⁻ D⁻)	(F⁻ 0)
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Control variables						
Prior problem-solving performance	1.30 (1.14)	1.63 (1.32)	1.60 (1.16)	1.53 (1.31)	1.72 (1.26)	1.41 (1.21)
Beliefs regarding the usefulness of educational evidence	3.74 (0.58)	3.76 (0.61)	3.85 (0.54)	3.85 (0.65)	3.74 (0.66)	3.89 (0.57)
Learning success						
Declarative EIRS	1.69 (1.14)	1.66 (1.19)	1.03 (0.93)	1.30 (1.12)	1.31 (0.91)	1.05 (0.91)
Near transfer problem-solving performance	2.33 (1.03)	2.55 (0.97)	2.01 (0.85)	2.10 (0.96)	2.16 (0.69)	1.91 (0.81)
Far transfer problem-solving performance	2.86 (1.02)	2.73 (1.10)	2.40 (1.00)	2.06 (0.96)	2.21 (0.92)	2.23 (0.92)
Learning process						
Error awareness	1.72 (0.60)	2.06 (0.50)	1.88 (0.48)	1.94 (0.55)	2.08 (0.54)	1.88 (0.51)

learning with the example-free instruction on dysfunctional procedures turned out to be more helpful in terms of estimating the own approach compared to learning without any instruction on dysfunctional procedures, with both forms of instruction on functional procedures (Tables 4, 6: Contrasts 3 and 6). The correlation analyses revealed significant positive correlations between error awareness and the three learning success variables with small to moderate effect sizes (Table 5).

3.4. Serial mediation (Research Question 3)

Two moderated serial mediation analyses were computed to examine to what extent the effect of the form of instruction on functional procedures (i.e., example-free vs. example-based) on near transfer problem-solving performance (Model 1) and far transfer problem-solving performance (Model 2) was serially mediated by *error awareness* as first stage mediator, and the *declarative EIRS* as second stage mediator for all three forms of instruction on dysfunctional procedures (Hypothesis 3; Figure 1). The indirect pathways are displayed in Table 8.

The participants' error awareness and declarative EIRS turned out to mediate the effect of the form of instruction on functional procedures on the students' near and far problem-solving performance (Model 1 and 2). In both models, this indirect effect was only significant for the conditions with the example-free instruction on dysfunctional procedures and without any instruction on dysfunctional procedures: when the instruction on functional procedures was supplemented by an example-free instruction on dysfunctional procedures or was not supplemented, the example-based instruction on functional procedures increased the students' error awareness. Consequently, the more the students were aware of the correctness of their own approach, the more developed was their declarative EIRS. The more developed the students' declarative EIRS was, the better they were able to solve near and far transfer problems.

It should be noted that the aforementioned effect of the form of instruction on functional procedures on near transfer problem-solving performance (Model 1) was completely mediated by error awareness and the declarative EIRS, as the direct effect did not reach the level of statistical significance (Table 8). The aforementioned effect of the form of instruction on functional procedures on far transfer problem-solving performance (Model 2) was also completely mediated in the condition without any instruction on dysfunctional procedures, and partially in the condition with the example-free instruction on dysfunctional procedures, as the direct effect was significant (Table 8).

TABLE 5 Bivariate Pearson correlations of control and dependent variables.

		1	2	3	4	5	6
1	Prior problem-solving performance	–					
2	Beliefs regarding the usefulness of educational evidence	0.04	–				
3	Declarative EIRS	0.07	0.13*	–			
4	Near transfer problem-solving performance	0.12*	0.14**	0.40**	–		
5	Far transfer problem-solving performance	−0.08	0.10*	0.20**	0.29**	–	
6	Error awareness	0.12*	0.15**	0.22**	0.34**	0.18**	–

$**p < 0.01$; $*p < 0.05$ (two-tailed).

TABLE 6 Results of planned contrast tests for dependent variables with significant main effect of the form of instruction on dysfunctional procedures.

Contrast	1		2		3		4		5		6	
Dependent variable	$F^+ D^+$ vs. $F^+ D^-$		$F^+ D^+$ vs. $F^+ 0$		$F^+ D^-$ vs. $F^+ 0$		$F^- D^+$ vs. $F^- D^-$		$F^- D^+$ vs. $F^- 0$		$F^- D^-$ vs. $F^- 0$	
	t	p	t	p	t	p	t	p	t	p	t	p
Declarative EIRS	0.11[a]	0.455	3.58[a]	**<0.001**	3.37[a]	**<0.001**	−0.09[a]	0.465	1.39[a]	0.083	1.66[a]	0.050
Near transfer problem-solving performance	−1.25[a]	0.107	1.97[a]	0.026[b]	3.40[a]	**<0.001**	−0.40[a]	0.344	1.18[a]	0.121	1.84[a]	0.034[b]
Error awareness	−3.65	**<0.001**	−1.70	0.045[b]	1.95	0.026[b]	−1.49	0.068	0.62	0.267	2.12	0.017[b]

p one-tailed. Adjusted level of significance by Bonferroni-Holm correction. Significant results are printed in bold letters.
[a]Corrected degrees of freedom due to violating homoscedasticity.
[b]Not significant due to Bonferroni-Holm correction.

Regarding the example-based instruction on dysfunctional procedures, none of the indirect paths reached statistical significance – neither in Model 1 nor in Model 2. Only the direct path from the form of instruction on functional procedures to far transfer problem-solving performance (Model 2) turned out to be significant (Table 8). The example-based instruction on functional procedures yielded better problem-solving performance without increasing error awareness and the declarative EIRS as long as an example-based instruction on dysfunctional procedures was also present.

4. Discussion

The aim of the present study was to examine the effects of a PS-I-based learning environment on pre-service teachers' ability to deal with problem situations in the classroom. The study focused on the effects of different forms of instruction after problem-solving in the domain of teacher education, with the aim of fostering a systematic, coherent, and evidence-informed reasoning approach. The study addressed evidence-informed reasoning not only from a perspective focusing on the use of scientific knowledge in practical situations, but especially from a perspective focusing on the cognitive processes of evidence-informed reasoning (Kiemer and Kollar, 2018; Csanadi et al., 2021). The internal validity of the study can be regarded as secured, as the six experimental conditions did not differ significantly with respect to possible confounding variables.

Regarding the students' learning success, the study focused on both basic declarative knowledge dimensions (i.e., declarative EIRS) and more complex procedural knowledge dimensions (i.e., near and far transfer problem-solving performance). Hypothesis 1 was partly

confirmed. As expected, the acquisition of the declarative EIRS as well as its application in the near and far transfer was supported by the example-based instruction on functional procedures. The findings suggest that comparing one's own solution approach with a specific exemplification of how to deal with a problem following the EIRS in a functional way contributes to the memorization of the script and helps to deal with problem situations in a more systematic, coherent, and evidence-informed way. The specific exemplification of the EIRS seems to promote pre-service teachers' ability to apply the EIRS not only to familiar, but even to new, unfamiliar situations. The results contribute to the extensive literature pointing to the benefits of example-based learning, and especially of worked examples (e.g., Renkl, 2017; van Gog and Rummel, 2018; Mayer, 2020).

However, the postulated superiority of providing an example-based instruction on dysfunctional procedures over providing an example-free instruction in terms of fostering pre-service teachers' learning success could not be confirmed. Regarding the development of a declarative EIRS and its application to a familiar problem situation (i.e., near transfer problem-solving performance), it was important for the students to learn with any kind of instruction on dysfunctional procedures. The analysis of typical dysfunctional procedures obviously helps students to internalize the EIRS and to apply it to a situation that is comparable to the training situation – regardless the specific form of instruction. The specific form of instruction on dysfunctional procedures does not seem to be relevant for the internalization of the EIRS, but particularly the fact that dysfunctional strategies of solving a problem situation are presented at all. Apparently, underlining typical dysfunctional solution approaches of other pre-service teachers leads attention to the principles to be learned (Große and Renkl, 2007; Oser et al., 2012). Surprisingly, when it came to applying the EIRS to

TABLE 7 Results of contrast tests for dependent variables with significant interaction effects.

Contrast	1		2		3		4		5	
Dependent variable	$F^+ D^+$ vs. $F^+ D^-$		$F^+ D^+$ vs. $F^+ 0$		$F^+ D^+$ vs. $F^- D^+$		$F^+ D^+$ vs. $F^- D^-$		$F^+ D^+$ vs. $F^- 0$	
	t	p	t	p	t	p	t	p	t	p
Far transfer problem-solving performance	0.74	0.230	2.67	**0.004**	4.59	**<0.001**	3.74	**<0.001**	3.65	**<0.001**

p one-tailed. Adjusted level of significance by Bonferroni-Holm correction. Significant results are printed in bold letters.

FIGURE 3
Interaction plot regarding the far transfer problem-solving performance.
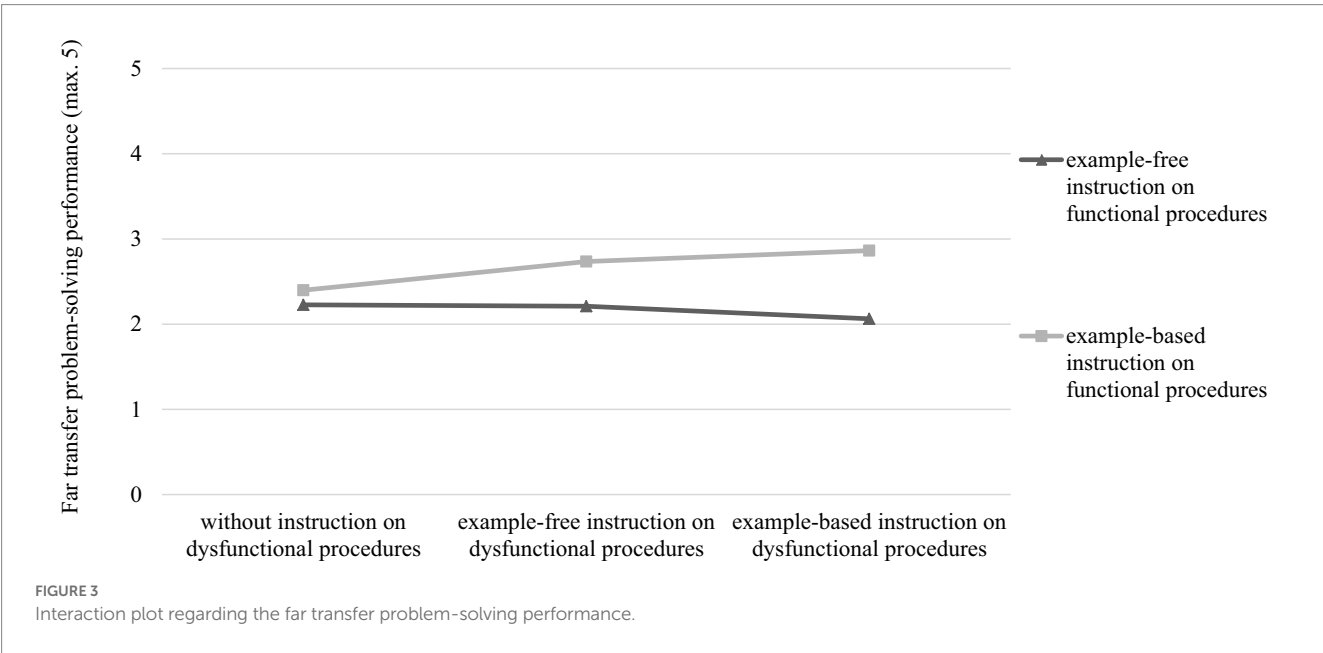
a novel, unfamiliar problem situation (i.e., far transfer problem-solving performance), instructing students on typical dysfunctional procedures did not directly foster their ability to solve the problem, but the specific combination of any kind of instruction on dysfunctional procedures with the example-based instruction on functional procedures did. The exemplification of a functional approach seems to play out its potential when it is complemented by typical dysfunctional approaches.

Error-based learning approaches are often criticized for the risk of learners internalizing dysfunctional procedures (Metcalfe, 2017). This critique cannot be supported by the findings of the present study. The contrary is true: the presentation of dysfunctional approaches seems to be crucial to promote pre-service teachers' problem-solving performance. Looking at the present findings on the benefits of instructions on dysfunctional procedures as a whole, acquiring knowledge about what not to do in principle when dealing with problem situations (i.e., *negative knowledge*; Oser et al., 2012) seems to help learners to apply the EIRS to other situations. It even seems necessary to provide both a specific illustration of a "good practice" approach and at least a description or illustration of a dysfunctional approach in the instruction phase of PS-I to enable the transfer of the EIRS to unfamiliar situations. In sum, the benefits of error-based learning (e.g., Große and Renkl, 2007; Loibl and Leuders, 2018, 2019) can be replicated by the present study.

The expectations set under Hypothesis 2 regarding error awareness was also partially confirmed. Surprisingly, although the ANOVA

results indicated that the form of instruction on functional procedures (i.e., example-free vs. example-based) had no impact on the quality of the student's assessments of their own approach, it must be noted that the mediation analyses hinted at this factor to be decisive for error awareness (Hypothesis 3). Considering the mediation results, it can be cautiously stated that pre-service teachers especially seem to benefit from instruction containing a specific illustration of a functional approach, in order to assess the correctness of their own approach.

Both the correlation analyses and the mediation analyses indicated that error awareness seems to be decisive for learning (Hypothesis 3), which is in line with the considerations on the key mechanisms of PS-I (Loibl et al., 2017; Nachtigall et al., 2020): Comparing one's own approach with an exemplification of a functional approach helps learners to become aware of their own errors, which then helps to further develop a mental schema of how to solve problems in principle. When dealing with problem situations, it seems important for pre-service teachers to have internalized the principle and operations behind the EIRS at a declarative level (Fischer et al., 2013), which in turn is promoted by comparing one's own approach with approaches of others. Therefore, the findings highlight the importance of estimating one's own approach in instruction after problem-solving for learning and correspond with findings on the benefits of comparisons for learning and transfer in general (e.g., Durkin and Rittle-Johnson, 2012; Loibl and Leuders, 2018, 2019).

However, the mechanism described above seems to apply particularly to instruction on functional procedures that is not supplemented by any

TABLE 8 Conditional direct and indirect effects for the moderated serial mediation models regarding near and far transfer problem-solving performance.

Tested paths	Model 1: Near transfer problem-solving performance						Model 2: Far transfer problem-solving performance					
	Without instruction on dysfunctional procedures		Example-free instruction on dysfunctional procedures		Example-based instruction on dysfunctional procedures		Without instruction on dysfunctional procedures		Example-free instruction on dysfunctional procedures		Example-based instruction on dysfunctional procedures	
	B	CI [LL; UL]	B	CI [LL; UL]	B	CI [LL; UL]	B	CI [LL; UL]	B	CI [LL; UL]	B	CI [LL; UL]
pv-ov	−0.038	[−0.264; 0.339]	0.257	[−0.046; 0.559]	0.118	[−0.244; 0.479]	0.112	[−0.216; 0.439]	**0.410***	**[0.057; 0.764]**	**0.720***	**[0.387; 1.053]**
pv-m1-ov	**0.068***	**[0.010; 0.144]**	**0.064***	**[0.009; 0.135]**	0.014	[−0.027; 0.064]	0.059	[−0.006; 0.151]	0.055	[−0.006; 0.146]	0.012	[−0.025; 0.062]
pv-m2-ov	−0.035	[−0.148; 0.071]	−0.087	[−0.033; 0.229]	0.122	[−0.003; 0.262]	−0.014	[−0.066; 0.033]	0.035	[−0.013; −0.104]	0.049	[−0.003; 0.125]
pv-m1-m2-ov	**0.030***	**[0.007; 0.063]**	**0.028***	**[0.005; 0.063]**	0.006	[−0.012; 0.027]	**0.012***	**[0.002; 0.028]**	**0.011***	**[0.001; 0.028]**	0.002	[−0.005; 0.013]

pv = predictor variable (factor: form of instruction on functional procedures); m1 = error awareness; m2 = declarative EIRS; ov = outcome variable (problem-solving performance). Mediating and dependent variables were z-standardized. Significance of regression coefficients (B) was interpreted referring to 95% confidence intervals (CI) not including zero. Significant results are printed in bold letters. *Significant result.

instructions on dysfunctional procedures or complemented only by example-free instructions on dysfunctional procedures. If there are no specific examples of typical dysfunctional procedures presented, it is important to have recognized one's own dysfunctional procedures, to build knowledge on how to solve and how not to solve problem situations in principle (i.e., *negative knowledge*; Oser et al., 2012). In contrast, when instruction on functional procedures is combined with a specific exemplification of dysfunctional procedures for a given problem, this knowledge seems to be acquired without the need to having recognized one's own errors before. This finding underlines the benefits of erroneous examples for learning success (e.g., Große and Renkl, 2007; Durkin and Rittle-Johnson, 2012; Barbieri and Booth, 2020). However, when it comes to estimating one's own approach, it might be more beneficial if the specific exemplifications of dysfunctional procedures match exactly with the dysfunctional procedures applied by the learners themselves (Loibl and Leuders, 2019); non-matching exemplifications might rather distract learners from recognizing own errors, which does not necessarily interfere with learning from the example itself.

4.1. Limitations and future directions

The present study is not without limitations. Concerning the methodology of the study, the sample size was too small to apply more refined statistical procedures such as structural equation modeling. The regression-based approach applied in the present study did not account for potential variable reciprocity or autoregressive effects. Since the study was considered as a short-term intervention, no claims can be made about long-term effects.

It is striking that even though all groups seem to have benefited from the PS-I approach, learning success was clearly below expectations; the mean values of the learning success variables proved to be far from the achievable maximum. The entire procedure, with the exception of the pre-test, was completed in one session to avoid

sample failures, forgetting effects and between-group communication. However, working on three scenarios in one session was perhaps too strenuous for some of the participants, which might have influenced their learning success negatively. The internalization of an EIRS might require more time and practice (Anderson, 1996). However, the collected log data does not provide any information about time-on-task and whether work was done on the tasks during the time the computer window was open or not. It should be noted that the benefits of the PS-I approach have been mainly shown for conceptual learning and transfer, but less for learning complex problem-solving skills and procedural knowledge facets (Loibl et al., 2017; Sinha and Kapur, 2019, 2021b). Since, in contrast to many other PS-I studies, the approach was not compared with another approach, such as direct instruction or instruction prior to problem-solving (Loibl et al., 2017; Sinha and Kapur, 2019, 2021b; Nachtigall et al., 2020), no conclusion can be drawn as to whether another approach would have been more appropriate to promote the acquisition and transfer of an EIRS.

Furthermore, the PS-I approach was not implemented in a "traditional" way (e.g., traditional productive failure; e.g., Kapur, 2012, 2014). In contrast to the fidelity criteria formulated by Sinha and Kapur (2021b) for PS-I and productive failure in particular, the students in the present study worked individually and online and had no social surround that might have influenced their learning process – partly for COVID-19-pandemic reasons and partly for testing reasons. Thus, students were neither able to work collaboratively nor were they advised and supported by a teacher during the instruction phase. An opportunity to work in groups and a social surround facilitation in both PS-I phases might have fostered the identification, elaboration, and organization of the critical features of the EIRS (Sinha and Kapur, 2021b). One further difference from traditional PS-I studies is that in the problem-solving phase of the present study the students did not generate totally intuitive solution ideas to the given problem (e.g., Kapur and Bielaczyc, 2012). They were provided with educational evidence and external guidance to use the steps of

evidence-informed reasoning, in order to reduce complexity. Therefore, the crucial mechanism of activating prior knowledge may have been mitigated in the training. However, the pre-test might have been more comparable to the traditional kind of problem-solving phase of the PS-I approach, which leads to another limitation of the study: the students might have also reflected on their procedure in the pre-test (on the influence of the presence and nature of pre-testing in PS-I see Sinha and Kapur, 2021b). But it should be noted that the pre-test was administered in a separate session 2 weeks before the training and that the problem scenario used in the pre-test was different from the problem scenario used in the intervention.

Another methodological limitation is that negative knowledge was not explicitly assessed. As one key goal of the present error-based approach is to make students aware of *how not to do something*, it would have been enlightening if statements could have been made about whether the students also remembered typical dysfunctional procedures or not (Loibl and Leuders, 2018).

Further, in previous PS-I studies, error awareness was usually measured at a subjective level using questionnaire items (e.g., Loibl and Rummel, 2014b; Glogger-Frey et al., 2015). In the present study, error awareness was operationalized by an objective measure to account for self-report biases (i.e., the task of comparing one's own solution to the given instruction). The instrument was conceptualized as a genuine part of the instruction phase of the implemented learning approach. Therefore, we are only able to draw conclusions about a rather specific form of error awareness that arises from the form of instruction depending on the experimental condition; no conclusions can be drawn about a rather *global awareness of knowledge gaps* that might have been triggered by cognitive processes in the problem-solving phase, without clarity about which specific component of the EIRS was lacking (Loibl and Rummel, 2014b). Since the written comparisons were at a rather superficial, principled level, it was not possible to analyze which specific errors the students became aware of. We cannot draw any inferences on the students' reflection quality, and whether they have really reflected on the rationale behind their dysfunctional procedures (Heemsoth and Heinze, 2016). Future studies should incorporate measures that allow precise conclusions on both global and specific awareness of knowledge gaps as well as the quality of reflections, in order to understand the key mechanisms of learning from instruction on functional and dysfunctional procedures after problem-solving.

The external validity might be another limitation of the present study. As the participants were already further advanced in their teacher studies, it is unclear whether novices would benefit in a similar way from both the treated topic and the learning approach or not (e.g., Kalyuga et al., 2001). In general, it needs to be clarified how much prior knowledge is needed to learn effectively from the PS-I approach incorporating typical dysfunctional procedures, and if rather adaptive forms of feedback are more efficient.

In addition to that, when there is a problem situation that a teacher needs to react immediately (such as a classroom disruption), the teacher has no time to work through the problem based on the EIRS. In time-pressured situations, teachers cannot search for appropriate evidence and reflect on the problem systematically (Renkl, 2022; Leuders et al., in press). Systematic, coherent, and evidence-informed reasoning based on the EIRS is more suitable for reflecting retrospectively on problem situations after teaching, and for making

considerations for further lessons in a scientific way. However, as the EIRS could become automated with time and experience, it might serve as heuristic to make short-term decisions in the classroom (such as giving immediate feedback in terms of Weiner's (1986) attribution theory). If teachers acquire additionally a "toolbox" (Renkl, 2022) of theories and empirical findings with time, they can solve problems in an evidence-informed manner, even if the time is limited.

A further limitation results from the fact that the summary of educational theories and findings that was provided in the training (and in the post-test) was precisely aligned with the problems described in the written case vignettes. In pedagogical practice, teachers are confronted with a multitude of information that they can apply in problematic situations. Therefore, ecological validity may be limited to that end, too. Future studies should consider addressing ecologically more valid problem-solving tasks (e.g., by means of authentic classroom videos) and providing multiple resources of information, containing for example also irrelevant, less suitable, and/or contradicting information.

In this context, it must be noted that the setting of the present study – at least implicitly – suggests that reasoning processes in which teachers apply educational knowledge are of higher quality than reasoning processes that are based on experiential knowledge. This position can be criticized as being too distant from practice (Wilkes and Stark, 2022). Preliminary confirmation of this critique can be found in empirical findings by Gegenfurtner et al. (2020) indicating that in-service teachers and school principals not only use evidence-based bodies of knowledge in their reasoning, but also episodic that is experience-based knowledge. Overall, research should broaden the theoretical perspective on evidence-informed reasoning in a way that is in line with assumptions on the development of (adaptive) expertise (Tsui, 2009; Bohle Carbonell et al., 2014) and focuses on the reflexive integration of both evidence-based and experience-based information.

4.2. Conclusion

The findings of the present study have strong implications not only for future research, but especially for teacher education practice, as they provide valuable indications for the design of learning environments that target the evidence-informed reasoning skills of pre-service teachers. Against the background of the calls for evidence-informed teaching practice, the study reveals that case-based learning approaches can support pre-service teachers in developing competences that might help in dealing with future problem situations in a systematic, coherent, and evidence-informed way. Although the present implementation of the PS-I approach differs from "traditional" PS-I studies conducted in STEM domains, it can be cautiously stated that the present case- and error-based approach with the sequence *instruction after problem-solving* works in the rather unstructured domain of teacher education – especially when examples of functional and dysfunctional procedures are included. In this context, the script perspective as outlined by Fischer et al. (2013), as well as Kiemer and Kollar (2018), has proven to be beneficial. The internalization of the EIRS and its transfer seem to particularly require concretization by illustrating how to proceed and how not to proceed when being confronted with various problem situations in the classroom. However, further research on the mechanisms of why and when

instruction on typical dysfunctional procedures works is needed. Motivational and attitudinal conditions should also be considered.

Overall, the findings underline the relevance of supporting future teachers in reflecting on problem situations by implementing learning opportunities in teacher education that not only focus on the acquisition of scientific knowledge (e.g., pedagogical-psychological theories and empirical findings), but also incorporate scaffolds for developing systematic evidence-informed problem-solving schemata. Future research should investigate how such problem-solving schemata can be further developed and consolidated, in order to be applied in everyday practice. Thus, research should also focus on the transition from learning with fictitious cases to learning with real-world situations in practice.

Data availability statement

The datasets presented in this article are not readily available because of legal reasons. Participants gave their consent to store, process and analyze the data as well as to the publication of the analysis results. Participants were assured that the data will not be distributed. Requests to access the datasets should be directed to TK-W, theresa.krause-wichmann@uni-saarland.de.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

Author contributions

TK-W and RS conceived of the presented idea and planned the experiment. TK-W and MG carried out the experiment. TK-W developed the theory and performed the computations. RS, MG, and IK verified the analytical methods. MG, CW, IK, and RS provided critical feedback and helped shape the research, analysis, and manuscript. TK-W took the lead in writing the manuscript with support and input from all co-authors. RS and IK supervised the project. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Anderson, J. R. (1996). ACT: a simple theory of complex cognition. *Am. Psychol.* 51, 355–365. doi: 10.1037/0003-066X.51.4.355

Atkinson, R. K., Derry, S. J., Renkl, A., and Wortham, D. (2000). Learning from examples: instructional principles from the worked examples research. *Rev. Educ. Res.* 70, 181–214. doi: 10.3102/00346543070002181

Barbieri, C. A., and Booth, J. L. (2020). Mistakes on display: incorrect examples refine equation solving and algebraic feature knowledge. *Appl. Cogn. Psychol.* 34, 862–878. doi: 10.1002/acp.3663

Bohle Carbonell, K., Stalmeijer, R. E., Könings, K. D., Segers, M., and van Merriënboer, J. J. (2014). How experts deal with novel situations: a review of adaptive expertise. *Educ. Res. Rev.* 12, 14–29. doi: 10.1016/j.edurev.2014.03.001

Brown, A. (1987). "Metacognition, executive control, self-regulation, and other more mysterious mechanisms" in *Metacognition, motivation, and understanding*. eds. F. E. Weinert and R. H. Kluwe (Hillsdale, NJ: Lawrence Erlbaum), 65–116.

Chase, C. C., and Klahr, D. (2017). Invention versus direct instruction: for some content, it's a tie. *J. Sci. Educ. Technol.* 26, 582–596. doi: 10.1007/s10956-017-9700-6

Chen, C. H., and Bradshaw, A. C. (2007). The effect of web-based question prompts on scaffolding knowledge integration and ill-structured problem solving. *J. Res. Technol. Educ.* 39, 359–375. doi: 10.1080/15391523.2007.10782487

Chi, M. T. H. (2000). "Self-explaining expository texts: the dual processes of generating inferences and repairing mental models" in *Advances in instructional psychology:*

Educational design and cognitive science. ed. R. Glaser, vol. *5* (Hillsdale, NJ: Lawrence Erlbaum Associates, Inc), 161–238.

Csanadi, A., Kollar, I., and Fischer, F. (2021). Pre-service teachers' evidence-based reasoning during pedagogical problem-solving: better together? *Eur. J. Psychol. Educ.* 36, 147–168. doi: 10.1007/s10212-020-00467-4

Dagenais, C., Lysenko, L., Abrami, P. C., Bernard, R. M., Ramde, J., and Janosz, M. (2012). Use of research-based information by school practitioners and determinants of use: a review of empirical research. *Evid. Policy* 8, 285–309. doi: 10.1332/174426412X654031

Durkin, K., and Rittle-Johnson, B. (2012). The effectiveness of using incorrect examples to support learning about decimal magnitude. *Learn. Instr.* 22, 206–214. doi: 10.1016/j.learninstruc.2011.11.001

Ferguson, L. E. (2021). Evidence-informed teaching and practice-informed research. *Z. Pädagog. Psychol.* 35, 199–208. doi: 10.1024/1010-0652/a000310

Fischer, F., Kollar, I., Stegmann, K., and Wecker, C. (2013). Toward a script theory of guidance in computer-supported collaborative learning. *Educ. Psychol.* 48, 56–66. doi: 10.1080/00461520.2012.748005

Fyfe, E. R., DeCaro, M. S., and Rittle-Johnson, B. (2014). An alternative time for telling: when conceptual instruction prior to problem solving improves mathematical knowledge. *Br. J. Educ. Psychol.* 84, 502–519. doi: 10.1111/bjep.12035

Ge, X., Chen, C.-H., and Davis, K. A. (2005). Scaffolding novice instructional designers' problem-solving processes using question prompts in a web-based

learning environment. *J. Educ. Comput. Res.* 33, 219–248. doi: 10.2190/5F6J-HHVF-2U2B-8T3G

Ge, X., and Land, S. M. (2003). Scaffolding students' problem-solving processes in an ill-structured task using question prompts and peer interactions. *Educ. Technol. Res. Dev.* 51, 21–38. doi: 10.1007/BF02504515

Ge, X., and Land, S. M. (2004). A conceptual framework for scaffolding ill-structured problem-solving processes using question prompts and peer interactions. *Educ. Technol. Res. Dev.* 52, 5–22. doi: 10.1007/BF02504836

Gegenfurtner, A., Lewalter, D., Lehtinen, E., Schmidt, M., and Gruber, H. (2020). Teacher expertise and professional vision: examining knowledge-based reasoning of pre-service teachers, in-service teachers, and school principals. *Front. Educ.* 5:59. doi: 10.3389/feduc.2020.00059

Glogger-Frey, I., Fleischer, C., Grüny, L., Kappich, J., and Renkl, A. (2015). Inventing a solution and studying a worked solution prepare differently for learning from direct instruction. *Learn. Instr.* 39, 72–87. doi: 10.1016/j.learninstruc.2015.05.001

Große, C. S., and Renkl, A. (2007). Finding and fixing errors in worked examples: can this foster learning outcomes? *Learn. Instr.* 17, 612–634. doi: 10.1016/j.learninstruc.2007.09.008

Hayes, A. F. (2021). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*, 3rd Edn. New York, NY: The Guilford Press.

Heemsoth, T., and Heinze, A. (2016). Secondary school students learning from reflections on the rationale behind self-made errors: a field experiment. *J. Exp. Educ.* 84, 98–118. doi: 10.1080/00220973.2014.963215

Helleve, I., Eide, L., and Ulvik, M. (2021). Case-based teacher education preparing for diagnostic judgement. *Eur. J. Teach. Educ.* 46, 50–66. doi: 10.1080/02619768.2021.1900112

Hetmanek, A., Wecker, C., Kiesewetter, J., Trempler, K., Fischer, M., Gräsel, C., et al. (2015). Wozu nutzen Lehrkräfte welche Ressourcen? Eine Interviewstudie zur Schnittstelle zwischen bildungswissenschaftlicher Forschung und professionellem Handeln im Bildungsbereich [for what do teachers use which kind of resources? An interview study on the interface between research and professional practice in education]. *Unterrichtswissenschaft* 43, 193–208.

Jacobson, M., Goldwater, M., Markauskaite, L., Lai, P. K., Kapur, M., Roberts, G., et al. (2020). Schema abstraction with productive failure and analogical comparison: learning designs for far across domain transfer. *Learn. Instr.* 65:101222. doi: 10.1016/j.learninstruc.2019.101222

Jonassen, D. H. (1997, 1997). Instructional design models for well-structured and III-structured problem-solving learning outcomes. *Educ. Technol. Res. Dev.* 45, 65–94. doi: 10.1007/BF02299613

Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educ. Technol. Res. Dev.* 48, 63–85. doi: 10.1007/BF02300500

Jonassen, D. H. (2011). *Learning to solve problems: A handbook for designing problem-solving learning environments*. New York: Routledge.

Joyce, K. E., and Cartwright, N. (2020). Bridging the gap between research and practice: predicting what will work locally. *Am. Educ. Res. J.* 57, 1045–1082. doi: 10.3102/0002831219866687

Kalyuga, S., Chandler, P., Tuovinen, J., and Sweller, J. (2001). When problem solving is superior to studying worked examples. *J. Educ. Psychol.* 93, 579–588. doi: 10.1037/0022-0663.93.3.579

Kapur, M. (2012). Productive failure in learning the concept of variance. *Instr. Sci.* 40, 651–672. doi: 10.1007/s11251-012-9209-6

Kapur, M. (2014). Productive failure in learning math. *Cogn. Sci.* 38, 1008–1022. doi: 10.1111/cogs.12107

Kapur, M., and Bielaczyc, K. (2012). Designing for productive failure. *J. Learn. Sci.* 21, 45–83. doi: 10.1080/10508406.2011.591717

Kiemer, K., and Kollar, I. (2018). "Evidence-based reasoning of pre-service teachers: a script perspective." in *Conference: International conference of the learning sciences: London, UK.*

Kiemer, K., and Kollar, I. (2021). Source selection and source use as a basis for evidence-informed teaching. *Z. Pädag. Psychol.* 35, 127–141. doi: 10.1024/1010-0652/a000302

Kirschner, P. A., Sweller, J., and Clark, R. E. (2006). Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educ. Psychol.* 41, 75–86. doi: 10.1207/s15326985ep4102_1

Kolodner, J. L. (1993). "Introduction" in *Case-based learning.* ed. J. L. Kolodner (Boston: Springer), 1–5.

Kramer, M., Förtsch, C., Seidel, T., and Neuhaus, B. J. (2021). Comparing two constructs for describing and analyzing teachers' diagnostic processes. *Stud. Educ. Eval.* 68:100973. doi: 10.1016/j.stueduc.2020.100973

Lange, C., Almusharraf, N., Koreshnikova, Y., and Costley, J. (2021). The effects of example-free instruction and worked examples on problem-solving. *Heliyon* 7:e07785. doi: 10.1016/j.heliyon.2021.e07785

Leuders, T., Wiedmann, M., and Loibl, K. (in press). "Evidenzorientierung in der Lehrkräftebildung. [Evidence-Based Teacher Education]" in *Evidenzbasierung in Lehrer*innenbildung. Schule und Bildungspolitik und -administration – Neue Befunde zu alten Problemen. [evidence-based teacher education. School and educational policy and administration – New findings on old problems].* eds. K. Besa, D. Demski, J. Gesang and H. Hinzke (Wiesbaden: Springer VS).

Loibl, K., and Leuders, T. (2018). Errors during exploration and consolidation – the effectiveness of productive failure as sequentially guided discovery learning. *J. Math.-Didakt.* 39, 69–96. doi: 10.1007/s13138-018-0130-7

Loibl, K., and Leuders, T. (2019). How to make failure productive: fostering learning from errors through elaboration prompts. *Learn. Instr.* 62, 1–10. doi: 10.1016/j.learninstruc.2019.03.002

Loibl, K., Roll, I., and Rummel, N. (2017). Towards a theory of when and how problem solving followed by instruction supports learning. *Educ. Psychol. Rev.* 29, 693–715. doi: 10.1007/s10648-016-9379-x

Loibl, K., and Rummel, N. (2014a). The impact of guidance during problem-solving prior to instruction on students' inventions and learning outcomes. *Instr. Sci.* 42, 305–326. doi: 10.1007/s11251-013-9282-5

Loibl, K., and Rummel, N. (2014b). Knowing what you don't know makes failure productive. *Learn. Instr.* 34, 74–85. doi: 10.1016/j.learninstruc.2014.08.004

Lysenko, L. V., Abrami, P. C., Bernard, R. M., and Dagenais, C. (2015). Research use in education: an online survey of Canadian teachers. *Brock Educ. J.* 25, 1–20. doi: 10.26522/brocked.v25i1.431

Lysenko, L. V., Abrami, P. C., Dagenais, C., and Janosz, M. (2014). Educational research in educational practice: predictors of use. *Can. J. Educ.* 12, 1–17. doi: 10.2190/mhur-4fc9-b187-t8h4

Mayer, R. E. (2020). Advances in designing instruction based on examples. *Appl. Cogn. Psychol.* 34, 912–915. doi: 10.1002/acp.3701

Metcalfe, J. (2017). Learning from errors. *Annu. Rev. Psychol.* 68, 465–489. doi: 10.1146/annurevpsych-010416-044022

Nachtigall, V., Serova, K., and Rummel, N. (2020). When failure fails to be productive: probing the effectiveness of productive failure for learning beyond STEM domains. *Instr. Sci.* 48, 651–697. doi: 10.1007/s11251-020-09525-2

Oser, F., Näpflin, C., Hofer, C., and Aerni, P. (2012). "Towards a theory of negative knowledge (NK): almost-mistakes as drivers of episodic memory amplification" in *Professional and practice-based learning.* eds. J. Bauer and C. Harteis, *Human Fallibility: The Ambiguity of Errors for Work and Learning*, vol. 6 (Dordrecht: Springer), 53–70.

Pekrun, R. (2006). The control-value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. *Educ. Psychol. Rev.* 18, 315–341. doi: 10.1007/s10648-006-9029-9

Piwowar, V., Barth, V. L., Ophardt, D., and Thiel, F. (2018). Evidence-based scripted videos on handling student misbehavior: the development and evaluation of video cases for teacher education. *Prof. Dev. Educ.* 44, 369–384. doi: 10.1080/19415257.2017.1316299

Posner, G. J., Strike, K. A., Hewson, P. W., and Gertzog, W. A. (1982). Accommodation of a scientific conception: towards a theory of conceptual change. *Sci. Educ.* 66, 211–227. doi: 10.1002/sce.3730660207

Renkl, A. (2002). Worked-out examples: instructional explanations support learning by self-explanations. *Learn. Instr.* 12, 529–556. doi: 10.1016/s0959-4752(01)00030-5

Renkl, A. (2017). "Instruction based on examples" in *Educational Psychology handbook series. Handbook of research on learning and instruction.* eds. R. E. Mayer and P. A. Alexander (New York, NY: Routledge), 325–348.

Renkl, A. (2022). Meta-analyses as a privileged information source for informing teachers' practice? A plea for theories as primus inter pares. *Zeitschrift für Pädagogische Psychol.* 36, 217–231. doi: 10.1024/1010-0652/a000345

Renkl, A., Mandl, H., and Gruber, H. (1996). Inert knowledge: analyses and remedies. *Educ. Psychol.* 31, 115–121. doi: 10.1207/s15326985ep3102_3

Roll, I., Holmes, N. G., Day, J., and Bonn, D. (2012). Evaluating metacognitive scaffolding in guided invention activities. *Instr. Sci.* 40, 691–710. doi: 10.1007/s11251-012-9208-7

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

Schank, R. C. (1999). *Dynamic memory revisited*. Cambridge, MA: Cambridge University Press.

Schank, R. C., and Abelson, R. P. (1977). *Scripts, plans, goals and understanding.* Hillsdale, NJ: Erlbaum.

Schwartz, D. L., and Bransford, J. D. (1998). A time for telling. *Cogn. Instr.* 16, 475–5223. doi: 10.1207/s1532690xci1604_4

Seidel, T., and Stürmer, K. (2014). Modeling and measuring the structure of professional vision in preservice teachers. *Am. Educ. Res. J.* 51, 739–771. doi: 10.3102/0002831214531321

Sinha, T., and Kapur, M. (2019). "When productive failure fails." in *41st annual conference of the cognitive science.* Montreal, Canada.

Sinha, T., and Kapur, M. (2021a). Robust effects of the efficacy of explicit failure-driven scaffolding in problem-solving prior to instruction: a replication and extension. *Learn. Instr.* 75:101488. doi: 10.1016/j.learninstruc.2021.101488

Sinha, T., and Kapur, M. (2021b). When problem solving followed by instruction works: evidence for productive failure. *Media Psychol.* 91, 761–798. doi: 10.3102/00346543211019105

Sinha, T., Kapur, M., West, R., Catasta, M., Hauswirth, M., and Trninic, D. (2020). Differential benefits of explicit failure-driven and success-driven scaffolding in problem-solving prior to instruction. *J. Ed. Psychol.* 113, 530–555. doi: 10.1037/edu0000483.?

Slavin, R. E., Cheung, A. C. K., and Zhuang, T. (2021). How could evidence-based reform advance education? *ECNU Rev. Educ.* 4, 7–24. doi: 10.1177/2096531120976060

Spensberger, F., Kollar, I., and Pankofer, S. (2021). Effects of worked examples and external scripts on fallacy recognition skills: a randomized controlled trial. *J. Soc. Work. Educ.* 58, 622–639. doi: 10.1080/10437797.2021.1963892

Thiel, F., Böhnke, A., Barth, V. L., and Ophardt, D. (2020). How to prepare preservice teachers to deal with disruptions in the classroom? Differential effects of learning with functional and dysfunctional video scenarios. *Prof. Dev. Educ.* 49, 108–122. doi: 10.1080/19415257.2020.1763433

Thomm, E., Sälzer, C., Prenzel, M., and Bauer, J. (2021). Predictors of teachers' appreciation of evidence-based practice and educational research findings. *Z. Pädagog. Psychol.* 35, 173–184. doi: 10.1024/1010-0652/a000301

Tsui, A. B. (2009). Distinctive qualities of expert teachers. *Teach. Teach.* 15, 421–439. doi: 10.1080/13540600903057179

van Gog, T., Paas, F., and van Merriënboer, J. J. G. (2004). Process-oriented worked examples. Improving transfer performance through enhanced understanding. *Instr. Sci.* 32, 83–98. doi: 10.1023/B:TRUC.0000021810.70784.b0

van Gog, T., and Rummel, N. (2010). Example-based learning: integrating cognitive and social-cognitive research perspectives. *Educ. Psychol. Rev.* 22, 155–174. doi: 10.1007/s10648-010-9134-7

van Gog, T., and Rummel, N. (2018). "Example-based learning" in *International handbook of the learning sciences*. eds. F. Fischer, C. E. Hmelo-Silver, S. R. Goldman and P. Reimann (New York, NY: Routledge), 201–209.

van Gog, T., Rummel, N., and Renkl, A. (2019). "Learning how to solve problems by studying examples" in *The Cambridge handbook of cognition and education*. eds.

J. Dunlosky and K. A. Rawson (Cambridge, Ma: Cambridge University Press), 183–208.

van Merriënboer, J. J. G., Kirschner, P. A., and Kester, L. (2003). Taking the load off a learner's mind: instructional design for complex learning. *Educ. Psychol.* 38, 5–13. doi: 10.1207/S15326985EP3801_2

van Merriënboer, J. J. G., and Sweller, J. (2010). Cognitive load theory in health professional education: design principles and strategies. *Med. Educ.* 44, 85–93. doi: 10.1111/j.1365-2923.2009.03498.x

Vosniadou, S. (2013). "Conceptual change in learning and instruction: the framework theory approach" in *The international handbook of conceptual change*. ed. S. Vosniadou. *2nd* ed (New York, NY: Routledge), 11–30.

Vosniadou, S. (2019). The development of students' understanding of science. *Front. Educ.* 4:32. doi: 10.3389/feduc.2019.00032

Wagner, K., Klein, M., Klopp, E., and Stark, R. (2016). Förderung anwendbaren bildungswissenschaftlichen Wissens anhand kollaborativem Lernen aus Fehlern [facilitation of applicable knowledge in teacher education by collaborative learning from instructional errors]. *Unterrichtswissenschaft* 44, 373–390.

Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer.

Westermann, K., and Rummel, N. (2012). Delaying instruction: evidence from a study in a university relearning setting. *Instr. Sci.* 40, 673–689. doi: 10.1007/s11251-012-9207-8

Wigfield, A., Tonks, S., and Klauda, S. L. (2016). "Expectancy-value theory" in *Handbook of motivation in school*. eds. K. R. Wentzel and D. B. Mielecpesnm. *2nd Edn*. (New York, NY: Routledge), 55–74.

Wilkes, T., and Stark, R. (2022). Probleme evidenzorientierter Unterrichtspraxis. Anregungen und Lösungsvorschläge. [problems of an evidence-oriented educational practice. Suggestions and solution approaches.]. *Unterrichtswissenschaft*. doi: 10.1007/s42010-022-00150-1

Wilkes, T., Stark, L., Trempler, K., and Stark, R. (2022). Contrastive video examples in teacher education: a matter of sequence and prompts. *Front. Educ.* 7:869664. doi: 10.3389/feduc.2022.869664

Yeh, C., and Santagata, R. (2015). Preservice teachers' learning to generate evidence-based hypotheses about the impact of mathematics teaching on learning. *J. Teach. Educ.* 66, 21–34. doi: 10.1177/0022487114549470