frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 01 November 2023
DOI 10.3389/feduc.2023.1257681

Check for updates

OPEN ACCESS

EDITED BY
Jennifer Paetsch,
University of Bamberg, Germany

REVIEWED BY
Nicolas Hübner,
University of Tübingen, Germany
Matthias Krepf,
University of Cologne, Germany
Daniel Scholl,
University of Siegen, Germany

*CORRESPONDENCE
Sonja Wedde
✉ sonja.wedde@uni-kassel.de

RECEIVED 12 July 2023
ACCEPTED 25 September 2023
PUBLISHED 01 November 2023

CITATION
Wedde S, Busse A, Bosse D and Gold B (2023) Comparing teaching examples: effects on the solution quality and learning outcomes of student teachers' professional vision of classroom management.
Front. Educ. 8:1257681.
doi: 10.3389/feduc.2023.1257681

COPYRIGHT
© 2023 Wedde, Busse, Bosse and Gold. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Comparing teaching examples: effects on the solution quality and learning outcomes of student teachers' professional vision of classroom management

Sonja Wedde[1]*, Annette Busse[1], Dorit Bosse[1] and Bernadette Gold[2]

[1]Institute of Educational Science, Faculty of Human Sciences, University of Kassel, Kassel, Germany, [2]Institute for Didactics and School Pedagogy, Faculty of Education, Psychology and Educational Research, Technical University Dortmund, Dortmund, Germany

Young teachers, in particular, often face difficulties in managing their classrooms. A professional vision of classroom management can support teachers in managing a classroom effectively. This element of professional competence is often promoted in teacher training using video-based approaches, e.g., videoing own lessons including expert feedback, analyzing instructional videos and accompanying transcripts and analyzing videos of own and others teaching. This study used the learning format of comparing contrasting cases to promote professional vision of classroom management among first-semester student teachers ($N = 127$). However, it diverged slightly from the norm in that its contrasting cases came from two auditive, not video, teaching examples. The effectiveness of comparing contrasting cases on the professional vision of classroom management was studied in two experimental groups. One group compared the contrasting cases using self-generated categories (invention activity, $n = 63$), and the other group compared the cases based on given categories (worked solution, $n = 64$). Invention activities and worked solutions have their origins in mathematics and science education and promise to activate prior knowledge and the curiosity of learners at the beginning of a new learning unit while at the same time highlighting deep features of a concept in order to enrich networked and organized knowledge. Professional vision of classroom management was assessed using a video-based online test in a pre-post design to determine whether it was more beneficial to compare contrasting cases with an invention activity or a worked solution. Additionally, using a coding scheme, we assessed the quality of the student teachers' task solutions and the relationship between the quality of the task solutions and the learning outcomes of the posttest. Both groups showed no difference between the benefits of the two learning formats. For the experimental group invention activity, we found negative correlations between the task solutions and the learning outcomes. We ascribe this correlation to a productive failure effect. We performed a second study ($N = 54$) with a follow-up test to investigate whether either of the two formats produced more sustainable results. Our studies showed that both experimental groups (invention activity: $n = 24$, worked solution: $n = 30$) also had a short-term increase in professional vision of classroom management. The invention activity, however, promoted the more sustainable acquisition of professional vision of classroom management. In conclusion, invention activities appear to be an appropriate task format for promoting professional vision of classroom management among student teachers. This finding is relevant for the

field of teacher education and improves insights into task formats that promote components of professional competence.

KEYWORDS

professional vision, preservice teacher education, classroom management, problem-solving prior to instruction, invention activities, worked solution, solution quality

1. Introduction

Classroom management has a significant impact on student learning outcomes (Wang et al., 1993; Seidel and Shavelson, 2007; Korpershoek et al., 2016) and improves learner motivation (Kunter and Voss, 2013). Yet, young teachers often have difficulty implementing classroom management strategies effectively in their teaching (Chaplain, 2008). This difficulty is partly because classroom management and the professional vision of classroom management have only recently become a major topic in teacher education research and teaching (e.g., Gold et al., 2021).

Classroom management, as one dimension of teaching quality, includes numerous strategies that promote students' academic and social–emotional learning (Evertson and Weinstein, 2006). To apply classroom management strategies successfully, teachers need the ability to perceive significant events in the classroom in a time-sensitive manner (Sherin, 2007; Seidel and Stürmer, 2014); they need a professional vision of classroom management in order to act in an appropriately adept way. Some studies have investigated the promotion of students' professional vision of classroom management via video-based approaches (Weber et al., 2018; Kramer et al., 2020; Gold et al., 2021). One common feature of these studies is the use of video-based training and pre- and posttests to assess professional vision of classroom management before and after training. However, novices tend to be overwhelmed by analyzing classroom videos (Erickson, 2007). Therefore, in our experimental studies, we present an innovative task format that aims to promote professional vision in student teachers by comparing auditive contrasting teaching examples. These auditive examples are intended to reduce the complexity of the analysis by focusing on the verbal level of a teacher's activities related to classroom management (Wedde et al., 2022). Thus, this exploratory study investigates the extent to which auditive cases can promote professional vision. Previous studies have focused on video-based or text-based formats. For example, a study by Syring et al. (2015) compared video cases with transcripts of those cases. Their study showed that working with videos caused higher extrinsic cognitive load among student teachers than working with only transcripts of the video.

Another aspect to consider when designing tasks for student teachers at their entry phase of a teacher training program is the multiple experiences they bring into their studies from their own time in school: student teachers already have preconceptions and beliefs about teaching quality (Calderhead and Robson, 1991). Therefore, it is important that learning formats activate prior knowledge and address these preconceptions and beliefs, as knowledge acquisition is influenced by them (Borko and Putnam, 1996).

To investigate what effect two different conditions, both working with auditive contrasting teaching examples, have on students' competence acquisition, we contrasted two experimental groups. In one experimental condition, the students compared the two teaching examples with self-generated categories (invention activity); in the other one, they compared the teaching examples using categories they had been given (worked solution). The objective of these studies was to evaluate which of the two task formats can positively influence student teachers' professional vision of classroom management.

2. Professional vision of classroom management

2.1. Classroom management

Classroom management functions as a support for instruction and requires the teacher to perform demanding actions (Martin and Sass, 2010). It can be categorized into three facets: monitoring, managing momentum and establishing rules and routines (Gold et al., 2021). Monitoring refers to Kounin's (1970) dimensions *withitness* and *overlapping*. Additionally, it includes strategies aimed at purposefully interrupting classroom disruptions with minimal distraction (Landrum and Kauffman, 2006; Simonsen et al., 2008; van Tartwijk, 2009). Managing momentum relates to the structuring of the lesson. It includes Kounin's (1970) dimensions *smoothness* and *momentum*. Similarly, for a lesson to be successful, it is important to establish a transparent teaching process as well as clarity about tasks, lesson objectives and behavioral expectations – establishing rules and routines (Emmer et al., 1994).

2.2. Professional vision

To be able to react appropriately to classroom events relevant to classroom management, teachers must be able to perceive these events in a time-appropriate manner (Seidel and Stürmer, 2014; Gold et al., 2021). Two interconnected sub-processes are key: the *noticing* of relevant teaching situations and the *knowledge-based reasoning* of these events in order to be able to react appropriately to the situation (Sherin, 2007; Seidel and Stürmer, 2014). Applying these sub-processes requires declarative, conceptual and case-based knowledge (Berliner, 2001; Stürmer et al., 2013; König et al., 2014). According to the *p*erception-*i*nterpretation-*d*ecision making model (PID model), professional vision is a situation-specific cognitive skill that mediates between dispositions, professional knowledge, affective-motivational characteristics and classroom performance (Blömeke et al., 2015; Blömeke and Kaiser, 2017). These cognitive processes occur before, during and after teaching in the classroom and thus are an integral part of teachers' professional actions. The promotion of professional

vision and its sub-processes, *noticing* and *knowledge-based reasoning*, thus should be one objective of teacher training programs and, recently, the number of studies on this topic has increased (Gold et al., 2021; König et al., 2022; e.g., Weber et al., 2018).

Research findings show that experts – in contrast to novices – can differentiate better between important and unimportant events in the classroom (Star et al., 2011; van den Bogert et al., 2014). Expert knowledge is more elaborate, more interconnected and more accessible than is the knowledge of novices (Borko and Livingston, 1989; Sabers et al., 1991). If disruptive behavior occurs, novices do not try to determine what might be triggering it and do not consider possible solutions to stop it – in contrast to experts. A study on this topic has been performed by Sabers et al. (1991), who evaluated novices', advanced learners' and experts' comments on classroom videos. They observed that, in contrast to experts, novice and advanced learners merely described visual and auditory cues and did not make connections between auditory and visual cues. Those participants made little overall reference to auditory aspects, using them only to describe what they saw. The novices and advanced learners were not able to interpret the spoken language and integrate the meaning of language to the whole observation. In contrast, experts included the teacher's spoken words and visual observations in their interpretations. To make meaningful interpretations about events in the classroom, experts link the observation of actions to the use of language and incorporate their professional knowledge (Sabers et al., 1991).

2.3. Promoting professional vision of classroom management

Classroom management is considered as a deep structure of teaching quality (Kunter and Voss, 2013). It is, thus, challenging for teachers to perceive events relevant to classroom management since these are on the level of learning processes and cannot be perceived on the surface level on the basis of features such as teaching methods. These challenges are evident in that novices differ from experts in their professional vision of classroom management (Wolff et al., 2015; König and Kramer, 2016; Gold and Holodynski, 2017; Stahnke and Blömeke, 2021). Thus, teacher education should promote the professional vision (of classroom management) of prospective teachers at an early stage through deliberate interventions.

To date, there have been various settings in which student teachers' professional vision of classroom management has been promoted using classroom videos and classroom transcripts (Weber et al., 2018; Kramer et al., 2020; Gold et al., 2021). Another approach is to compare teaching examples in five steps that can be used specifically to promote the sub-processes of professional vision [describing, interpreting, evaluating, and explaining (Sherin and van Es, 2009; Seidel and Stürmer, 2014)]: description, classification into categories, juxtaposition, summarization, and conclusion (Wedde et al., under review).

In our studies, auditive teaching examples of classroom management were used as contrasting cases for the comparison approach. In our lesson examples, excerpts from classroom scenes were listened to. Each of them contained audio track sequences of a classroom video. The first auditive teaching example represented a less successful, the second auditive teaching example a more successful implementation of classroom management strategies by the teacher. Listening to auditive teaching examples has the advantage of the participants being able to focus on verbal strategies of classroom management. These tasks overcome, for example, the issues expressed in the study by Sabers et al. (1991): novices included the auditive level of videos only descriptively and not analytically in their annotation. Practicing lesson analysis via auditive examples can be an effective way to bring the relevance of language into the focus of learners' analysis.

3. Comparison in the problem-solving prior to instruction approach

The problem-solving prior to instruction approach provides an effective method of comparing teaching examples, making it a very effective way for learners to critically analyze classroom management strategies. This approach is characterized by two steps: first, learners start in the problem-solving phase and subsequently receive instruction in the following phase (Loibl et al., 2017). The purpose of this approach is to activate learners' prior knowledge and to allow them to develop networked and organized knowledge. While working on the task, learners are expected to become aware of their own knowledge gaps, focus on deep features of the topic to be learned and develop curiosity regarding the new topic (Glogger-Frey et al., 2015; Loibl et al., 2017). Comparison as a learning method that can be used within this approach has been widely researched in the school context. Those studies are especially based in the mathematical and scientific domains (Schwartz et al., 2011; Loibl and Rummel, 2014b). Several meta-analyses found comparing to be effective for learning in both schools and higher education (Marzano et al., 2001; Apthorp, 2010; Alfieri et al., 2013).

3.1. Invention activities and worked solutions

Two promising learning formats following the problem-solving prior to instruction approach are invention activities and worked solutions. Comparisons are integrated into these learning formats by using contrasting cases. In the first step of *invention activities*, learners are given a set of contrasting cases with a prompt to develop a solution to the given problem by comparing the cases. For the *worked solution*, learners are generally presented with a sample solution for the contrasting cases they need to study as they work through the task (Renkl, 2014). In the subsequent instructional phase, the learners are presented with the canonical solution to the task and instruction is given on the topic to be learned (Schwartz et al., 2011).

In a comparison task, learners need to identify, explain and organize features. A comparison of constructed cases helps highlight the deep features of the concept to be learned (Wedde et al., 2022). Recognizing deep features is a significant part of working through the learning formats. It is also known from research on expertise that experts tend to categorize principles by deep features; novices tend to categorize through surface features (Chi et al., 1981). If the new knowledge of the concept to be learned is stored via deep features, it can be applied better to new situations, and flexible knowledge is developed (Kapur and Bielaczyc, 2012; Loibl et al., 2017). The auditive contrasting cases in this study aim to reduce the complexity of teaching analysis, to make essential strategies of classroom

management identifiable through contrasting and to provide negative knowledge about classroom management, i.e., knowledge about non-meaningful pedagogical practices (Oser et al., 2012; Wedde et al., 2022).

Comparing as an integrated task in invention activities can be more effective than simply asking students to find similarities and differences (Chi et al., 2012; Chin et al., 2016). Many studies have shown that invention activities promote the acquisition of conceptual rather than procedural knowledge (Loibl and Rummel, 2014a,b; Loibl et al., 2017; Weaver et al., 2018). Overall, studies contrasting the worked solution with the invention activity showed that the worked solution tended to be superior. The worked solution is considered the preferable format, especially for learners with little expertise and in the context of cognitive load theory, when it comes to task formats with a high level of element interactivity (Sweller et al., 1998). Element interactivity depends not only on the expertise of learners but also on the complexity of the learning material and refers to the elements that learners have to process in working memory to find a solution (Ashman et al., 2020).

In contrast to the worked solution effect, the generation effect, related to invention activity, assumes that learners who have to develop something while working on a task acquire knowledge more sustainably (Bertsch et al., 2007). In their study, Chen et al. (2015) found that the generation effect can occur with low element interactivity and that learners achieved better results when they had to develop something themselves. In contrast, when element interactivity was high, learners achieved better learning outcomes when they worked with a worked solution. This finding indicates that task formats with a low level of element interactivity, which contain prompts to generate something, promote learning.

Studies on the impact of when instruction is given (problem-solving prior to instruction or *vice-versa*) and on using contrasting cases have shown mixed results. One study (Loibl et al., 2020) could not find any effects: There were no pre- and posttest differences in conceptual understanding between the experimental groups that completed the task either before or after instruction and the task with or without contrasting cases. The authors of that study, thus, highlight that it is unclear which condition is beneficial to learning. They argue that novices, in particular, often did not use the beneficial features of the contrasting cases and may need additional guidance to do so. Learners who work with contrasting cases without additional guidance may tend to discover surface rather than deep features of the concept being learned; they would then perform low in the posttest. However, comparing contrasting cases has been shown to help learners identify more deep features of a concept (Loibl and Rummel, 2014a). Roelle and Berthold (2015) found that the experimental group with additional support for comparing contrasting cases, similar to a worked solution, performed better on the posttest in terms of conceptual knowledge than the other two experimental groups that did not receive additional support for comparing or did not complete a preparatory task. A study involving student teachers on the topic of learning strategies showed that the group working on the worked solution performed better regarding conceptual understanding on the posttest, although in one of these studies, for example, the invention activity resulted in significantly more knowledge gaps after task completion (Glogger-Frey et al., 2015). Those authors concluded that, in domains such as educational studies, working with a worked solution would be particularly suitable for knowledge acquisition and transfer and that long-term knowledge acquisition would have to be checked again in further studies (Glogger-Frey et al., 2015, 2022).

3.2. Relation between solution quality and learning outcomes

Previous studies have evaluated different aspects of learner solutions from the problem-solving phase regarding invention activities and worked solutions. We summarize the findings of these studies using the term solution quality. These findings lead to the question as to how solution quality is related to learning outcomes. There are different approaches to addressing this question. In one approach, the assumption is that a high solution quality results in advantages for the learning process since the problem-solving phase can support learners in discovering deep features of a concept to be learned (Loibl et al., 2017). This support is beneficial for the learning process because the deeper features are identified during the problem-solving phase, and the learners can focus on the less deep features during the instruction phase (Roll et al., 2011; Loibl and Rummel, 2014a). Alternatively, other researchers assume that failure during the problem-solving phase is beneficial for the learning process. This approach is called productive failure (Kapur and Bielaczyc, 2012). Important in productive failure is that learners become aware of their knowledge gaps and that, during instruction, they receive an explanation as to why a solution does not work (Sinha et al., 2021).

Previous studies also show a mixed picture of how solution quality relates to learning outcomes. In their study with psychology students, Wiedmann et al. (2012) found that high solution quality in the solution attempts correlated positively with procedural and conceptual knowledge in a follow-up quiz. In their experiment with student teachers on learning strategies, Glogger-Frey et al. (2015) only assessed the correctness of the solution attempts of the experimental group invention activity. The results also showed a positive correlation between the solution quality and posttest scores. In another study with student teachers, Glogger-Frey et al. (2022) found no correlations between the number of solution attempts or the number of optimal or suboptimal solution attempts for the experimental group invention activity. A further study also evaluated solutions only for the experimental condition invention activity. The study found that the quality of the solutions correlated positively with the results in the posttest (Roelle and Berthold, 2015). In their study on the problem-solving prior to instruction approach, Loibl and Rummel (2014a) found that, although the solutions of the group working with contrasting cases were of higher quality, the quality was not related to conceptual knowledge in the posttest. A recent study conducted by Sinha et al. (2021) with students in a course on data science yielded results in line with the productive failure approach: Learners who worked with a failure-driven task (learners were prompted toward an incorrect solution) during the problem-solving phase achieved higher conceptual understanding in the posttest than learners who worked with a success-driven task (learners who were prompted toward a correct solution). Hence, solving a task less successfully may well lead to better learning outcomes than solving a task more successfully.

The aforementioned studies, which are representative of research on the relationship between solution quality and learning outcome, show a mixed picture (Sinha et al., 2021). The findings suggest either a

positive relationship between solution quality and learning outcomes (Wiedmann et al., 2012; Glogger-Frey et al., 2015; Roelle and Berthold, 2015), a low solution quality and better learning outcomes (Sinha et al., 2021), or no correlations between solution quality and learning outcomes (Loibl and Rummel, 2014a). Moreover, it remains to be clarified to what extent these results can be replicated for educational research. The studies mentioned above were mostly conducted in the scientific and mathematical fields. Similarly, limited studies have shown that the worked solution is the more effective task format for educational studies (Glogger-Frey et al., 2022). The results of our studies up to now indicate that, at the process level, the worked solution task format is more appropriate (Wedde et al., 2022, under review).

4. Objectives

This theoretical and empirical framework shows that there has been no research on invention activities and worked solution regarding the potential of these task formats to promote professional vision, a situation-specific cognitive skill (Blömeke et al., 2015; Blömeke and Kaiser, 2017), among student teachers. As already outlined above, there is inconsistent evidence on whether the invention activity or the worked solution is the preferable learning format for novices – this preferability is also significantly related to the learning object, the learning goal and the element interactivity. For our learning environment, we assume a high element interactivity: The students had to recognize different, presumably unfamiliar, dimensions of classroom management and integrate them into the comparison, in addition to simultaneously performing cognitive processes. In previous studies, conceptual knowledge and conceptual understanding was often assessed in the posttest (Wiedmann et al., 2012; Loibl and Rummel, 2014a,b; Glogger-Frey et al., 2015; Roelle and Berthold, 2015; Weaver et al., 2018; Loibl et al., 2020; Sinha et al., 2021; Glogger-Frey et al., 2022). Professional vision relates to this because it is composed of knowledge-based sub-processes, which are based on declarative and conceptual knowledge (Stürmer et al., 2013; König et al., 2014).

In addition, there are different study results on the question to what extent learners' solution quality is related to their learning outcomes for task formats in problem-solving prior to instruction approach (Wiedmann et al., 2012; Loibl and Rummel, 2014a; Glogger-Frey et al., 2015; Roelle and Berthold, 2015; Sinha et al., 2021; Glogger-Frey et al., 2022). The present studies will contribute to this research by investigating whether there is a positive or negative relationship between solution quality and learning outcomes, as well as the question of which of the learning formats can initiate a situation-specific cognitive skill in the long term.

Through the task format of comparing two teaching examples, case-based knowledge is initiated. Additionally, the comparison and the corresponding instruction on the topic of classroom management are designed to promote declarative and conceptual knowledge on this topic. Thus, the competence to analyze teaching events without the urge to interact could be developed. By means of a video-based online test (Gold and Holodynski, 2017), which was used in the pre- and posttest for Study 1 (see Section 5) and also in the follow-up in Study 2 (see Section 6), students had to transfer this competence to videos in which auditive and visual elements are combined.

Consequently, the research questions of these experimental studies are

> *RQ1*: To what extent does comparing two contrasting auditive teaching examples promote novices' professional vision of classroom management when comparing (a) with self-generated categories (invention activity) or (b) with given categories (worked solution)? (see Study 1, Section 5 and Study 2, Section 6)
>
> *RQ2*: Are students' learning process outcomes (awareness of knowledge gaps, analytical and content-related solution quality) related to the development of their professional vision of classroom management? (see Study 1, Section 5)
>
> *RQ3*: To what extent can comparing teaching examples promote a long-term skill acquisition regarding professional vision of classroom management using the two different task formats? (see Study 2, Section 6)

5. Study 1

5.1. Hypotheses

Study 1 examined the extent to which the learning format can initiate professional vision of classroom management in the short term by comparing the two experimental conditions and investigating the extent to which there is a positive or negative correlation between solution quality and learning outcomes. Teacher education studies on invention activities and worked solution with a high element interactivity found that the worked solution condition is more suitable (Glogger-Frey et al., 2015, 2022). Other study results also proved that, for learning material with a high element interactivity, the worked solution might be the preferable format (Chen et al., 2015). For RQ1, we formulated the following hypothesis.

> *H1*: The experimental group worked solution (WS) achieves higher scores in the posttest than the experimental group invention activity (IA) after controlling for pretest differences.

Overall, the problem-solving phase is assumed to be related to the posttest. Therefore, we included the variables of analytical and content-related solution quality in this study (see Section 5.2.4 for a description of these variables; Wedde et al., 2022, under review). However, it remains to be clarified whether these correlations between the variables of analytical and content-related solution quality are positive or negative. Negative correlations would speak for a productive failure effect (Kapur and Bielaczyc, 2012). Therefore, we also formulated non-directional hypotheses to help answer RQ2.

> *H2a*: The IA experimental group perceives more knowledge gaps after working on the task than the WS experimental group after controlling for pretest differences.
>
> *H2b*: The more knowledge gaps were perceived after completing the task, the better students performed on the posttest.
>
> *H3a*: The number of perceived surface features is related to the posttest scores.

> *H3b*: The number of perceived deep features is related to the posttest scores.
>
> *H3c*: The number of mentioned categories is related to the posttest scores.
>
> *H3d*: The depth of comparison is related to the posttest scores.

5.2. Method

5.2.1. Sample

The overall sample consisted of 145 student teachers in their first semester of their studies at the University of Kassel, Germany. Seventy-four were randomly assigned to the IA experimental group (67.6% female; age: $M = 22.1$, $SD = 4.7$). The other 71 students were randomly assigned to the WS experimental group (66.2% female; age: $M = 21.4$, $SD = 4.8$). All students were enrolled in a teacher training program for secondary schools.

For the 145 participants, the time watching the videos was noted. A video filter was used to remove any cases with a deviant video time from the analysis. A deviant video time is characterized by the fact that all four videos to be analyzed were watched much shorter or much longer than the actual length of the video. Thus, the video filter removed cases that demonstrated deviant video times at pre- or post-test. The video filter resulted in the overall sample being reduced to 127 participants (IA: $n = 63$, WS: $n = 64$), a sample dropout of 12.4%. A similar number of cases were removed from both experimental groups in this process, χ^2 (1) = 0.84, $p = 0.36$. The cases with a deviant video time demonstrated higher professional vision of classroom management in the pretest than the cases that demonstrated a compliant video time (see Table 1). The groups did not differ regarding age, $F(1,142) = 0.004$, $p = 0.95$, $\eta^2 = 0.00$, and regarding gender, χ^2 (1) = 1.19, $p = 0.28$. The analytical and content-related solution quality were only evaluated for the cases that were also included in the data evaluation of the pretest and posttest (Wedde et al., 2022, under review).

5.2.2. Research design

The experimental study was conducted in the lecture "Introduction to the pedagogy of secondary schools" for first-year students in educational science, which was held by a professor. This lecture gives an overview of significant topics in educational science besides classroom management, including inclusive education or teaching quality. After the introductory session, the students completed the pretest. In the following session, the students worked on the task comparing the auditive teaching examples. Working on the task took about 45 min. After completing the task, awareness of knowledge gaps was assessed using a standardized questionnaire. In the following session, the students received instruction on classroom management and were also presented with the canonical solution to the task. One week later, the posttest was conducted. Students completed the pretest and posttest during course time. Due to the COVID-19 pandemic, the lecture was conducted online via videoconference. See Figure 1 for an overview of the research design.

TABLE 1 Differences between groups included or excluded by the video filter.

Variable	*M* (*SD*)		*U*	*z*	*p*
	Through video filter ...				
	... excluded $n = 18$	... included $n = 127$			
Pretest	0.47 (0.18)	0.34 (0.18)	677.5	−2.79	<0.01

A coefficient of agreement from 0 to 1 is used for the pretest and posttest. Since the variable is not normally distributed (Shapiro–Wilk test, $p < 0.05$), a Mann-Whitney-U test was calculated to determine differences between the groups.

5.2.3. Treatment

The goal of this treatment is not only promoting the student teachers' professional vision of classroom management but also raising curiosity and interest in learning about classroom management. The examples were the audio tracks sequences of two constructed classroom videos. These examples were presented as podcasts for the students. In the first teaching example, strategies of classroom management were used less successfully; in the second teaching example, they were used successfully (Wedde et al., 2022). For the first lesson example, i.e., the less successful example, the teacher has not established any rules, does not manage to establish clear transitions between the lesson phases, or frequently focuses on the learners' misbehavior. In contrast, in the second example, students listen to how the teacher has established and successfully applies rules, has structured the lessons in a meaningful way, as well as how her transitions are clear and how she praises the learners' behavior. The order of the two classroom examples was intentional: By listening to how classroom management strategies are less successfully applied, students can initiate negative knowledge about strategies for effective classroom management. Thus, learners are exposed to less effective strategies before they are exposed to effective strategies. This contrast with a less successful classroom example not only highlights the effective strategies but also emphasizes the importance of these strategies, some of which seem self-evident and trivial (Oser et al., 2012). The students listened to these examples at the beginning of the task. Before the IA experimental group could compare the two teaching examples, students were asked to develop categories for comparing the two contrasting cases. The WS experimental group was given categories that they had to use to identify the differences between the two teaching examples (i.e., the given categories were managing transitions, rules, routines, communication by the teacher and managing disruptions).

The chosen task formats invention activity and worked solution began with the problem-solving phase, in which students were asked to compare contrasting cases by generating own categories (IA) or with given categories (WS). Beginning with this assignment was meant to raise student teachers' epistemic curiosity and situational interest in classroom management, and students were meant to become aware of their knowledge gaps. Additionally, prior knowledge was activated (Glogger-Frey et al., 2015; Loibl et al., 2017). Thus, student teachers should not only become interested in learning about classroom management, but they should also become aware of what they do not know about this concept. Another goal of completing this assignment was for students to acquire case-based knowledge (Berliner, 2001) as well as negative knowledge (Oser et al., 2012) about classroom management. By comparing contrasting cases, effective strategies of classroom management should become salient, and students should learn how to apply classroom management strategies

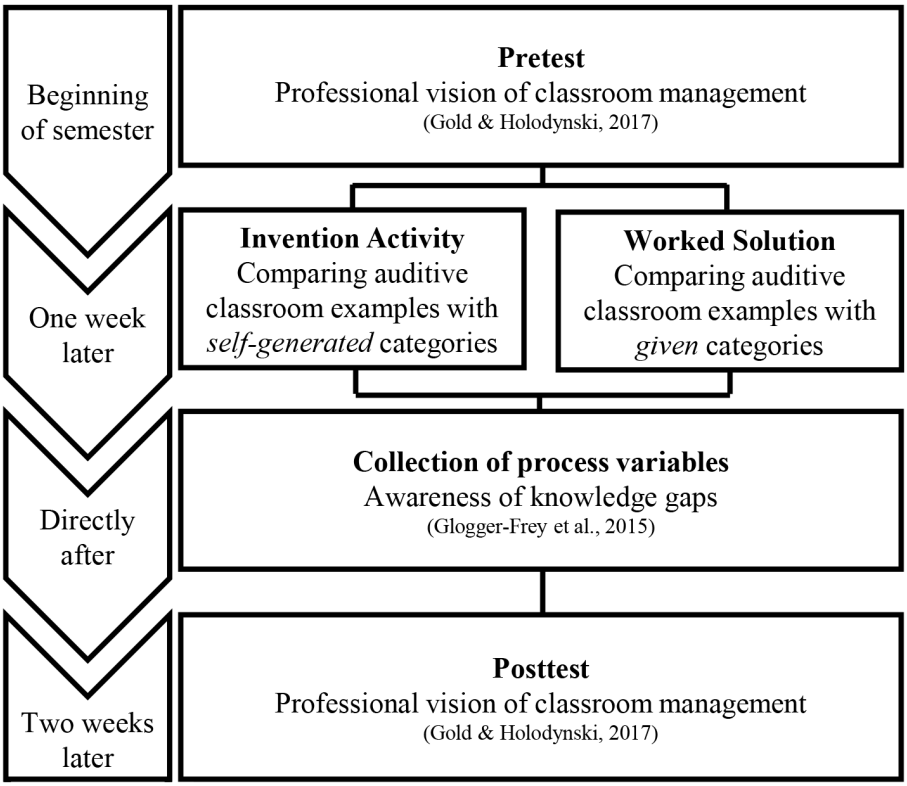

FIGURE 1
Research design of study 1.

effectively in the classroom (Wedde et al., 2022). During instruction, this case-based knowledge was enriched by declarative as well as conceptual knowledge.

The auditive teaching examples set the focus on the teacher, with student teachers being introduced to the role of the teacher. Often student teachers tend to focus on irrelevant features of the presented classroom example in their analysis (Sabers et al., 1991; Star et al., 2011). Through the auditive format, the students in both teaching examples were not in the spotlight, which could make it easier for student teachers to focus on the strategies applied by the teacher. One disadvantage of auditive teaching examples may be the inaudible non-verbal classroom management strategies of the teacher. This part of the analysis will be introduced to the student teachers later in their studies so that the analysis of auditive teaching cases serves as an introduction to analyzing classroom scenes. According to the modality principle of the cognitive theory of multimedia learning, it is more efficient to present information using both visual and auditory channel (Low and Sweller, 2009; Mayer, 2009). This argument may not be fully applicable to student teachers who acquire analytical skills at the beginning of their studies. Novices cannot cognitively process the large amount of information presented in videos at the same rate as it is presented and can quickly become cognitively overloaded while analyzing videos (Erickson, 2007). According to our studies, after comparing the auditive teaching examples, the student teachers experienced relatively low extrinsic and intrinsic cognitive load (Wedde et al., 2021, 2022, under review). The results of this study will provide insight into the extent to which auditive cases can enhance professional vision.

5.2.4. Instruments

A standardized video-based online test was used for the pre- and posttest to assess professional vision of classroom management. This variable indicates a coefficient of agreement with an expert rating from 0 to 1 (Gold and Holodynski, 2017). Awareness of knowledge gaps, adapted from Glogger-Frey et al. (2015), was assessed directly after completing the task by a standardized questionnaire using eight items on a scale from 1 "strongly disagree" to 6 "strongly agree" ($\alpha = 0.84$, e.g., "My knowledge was insufficient to solve the task."). Besides the awareness of knowledge gaps, the analytical and content-related solution quality were also used to check whether there were any relationships between learning processes outcomes and the acquired professional vision of classroom management.

To assess the analytical and the content-related solution quality of the tasks, we used two coding measures that had been developed during previous studies (Wedde et al., 2022, under review). The analytical solution quality means "the quality of the learning processes induced by the task of comparing contrasting cases" (Wedde et al., under review, p. 16). For the analytical solution quality, five levels of effective comparison were used in the depth of comparison scale: Description, Classification into categories, Juxtaposition, Summarization, Conclusion (Wedde et al., under review). If one of the steps was completed, one point was given. If one step was not achieved, zero points were given. Before a higher step could be reached, the previous step must have been fulfilled. A total of five points per task solution could be achieved for the five steps (Wedde et al., under review). A sum score was created on the given points (Plöger et al., 2020). See Table 2 for an overview of the categories.

TABLE 2 Categories of the analytical solution quality.

Step		Condition	Definition
1	Description		Reproduction of classroom situations and/or actions relevant to classroom management
2	Classification into categories	IA 2a	References to classroom situations and/or actions from both examples based on self-generated classroom management categories for comparison
		WS 2b	Collection of specific and relevant classroom management actions and/or situations from both examples according to the given categories
3	Juxtaposition		Comparison of both examples using the self-generated or given categories
4	Summarization		Development of summaries based on the comparative analysis of which of the two examples represented a better use of classroom management strategies
5	Conclusion[a]		Development of conclusions on the comparative analysis of learning effectiveness in both teaching examples and benefits to students' future teaching practice

Step 2 is divided into two steps, each relating to the two experimental conditions. Adapted from Wedde et al., under review, p.18. [a]*Conclusion* was not included in the scale as this step was not found in the students' solutions.

To assess the content-related solution quality, we used surface and deep features for the three facets of classroom management (monitoring, rules & routines, and lesson structure) to be recognized in the two auditive contrasting cases that were based on the state of research on classroom management. For each facet, both the surface and the deep features were combined into one variable and included in the analysis. Surface features were defined as behavioral and well-audible features, such as aspects of teacher communication, like appearing to be annoyed. In contrast, deep features refer to aspects of the teaching-learning process or also, for example, to how the teacher manages disruptions. The evaluation additionally included how many categories were mentioned in each solution (Wedde et al., 2022). See Table 3 for selected sample categories of the classroom management facet monitoring.

Both coding systems were coded independently by two coders for all task solutions. Due to the heterogeneity of the solutions, consensual coding was used. During regular meetings, agreement was reached on values that did not match. Thus, this approach can also be considered reliable (Guest et al., 2011; see Supplementary material for examples and explanations of the coding).

5.2.5. Data analysis

Means and standard deviations were calculated for all data. To examine the extent to which the two experimental groups differed at posttest and regarding the awareness of knowledge gaps, analyses of covariance (ANCOVAs) were calculated. The posttest and awareness of knowledge gaps served each as the dependent variable and the experimental groups as the independent variable. Although there was no significant difference between the IA and WS experimental groups in the pretest, $F(1,125) = 3.18$, $p = 0.08$, $\eta^2 = 0.03$, the pretest was included as a control variable to increase statistical power for all analyses. Before the ANCOVAs were calculated, the prerequisites were checked. For the individual variables, a normal distribution can be assumed for both experimental groups. The variance of the variable posttest ($p = 0.89$) is homogeneous. For the variable awareness of knowledge gaps, the variance of the variable is not homogenous ($p < 0.05$). Since both experimental groups were almost the same size, the ANCOVA could still be calculated. Moreover, Pearson correlations were calculated between the solution quality variables and the posttest score. SPSS 27 was used for data analysis.

An *a priori* power analysis was conducted using G*Power version 3.1.9.7 (Faul et al., 2007) for sample size estimation for an ANCOVA. A medium effect size was assumed for the power calculation, as no comparable studies were available. After calculation, $N = 128$ participants would have been required with a significance level of $\alpha = 0.05$ and a power $(1\text{-}\beta) = 0.80$ for a medium effect size. The sample size of $N = 127$ is only slightly smaller, resulting in a similar post-hoc power of $(1\text{-}\beta) = 0.80$.

5.3. Results

5.3.1. Learning outcome (H1)

Table 4 shows the means and standard deviations for the pre- and posttest. The overall sample as well as the two subsamples IA and WS improved from the pretest to the posttest. The two experimental groups barely differed in the posttest and developed almost equally from the pretest to the posttest. The ANCOVA showed no significant effect for the experimental group at posttest, $F(1,124) = 1.66$, $p < 0.20$, $\eta^2 = 0.01$, but a highly significant effect of pretest, $F(1,124) = 87.66$, $p < 0.001$, $\eta^2 = 0.41$. Thus, hypothesis H1 must be rejected.

5.3.2. Relationship between solution quality and learning outcomes (H2 & H3)

There was a small difference between experimental groups IA ($M = 3.01$, $SD = 0.91$) and WS ($M = 2.78$, $SD = 0.75$) for the awareness of knowledge gaps after completing the task; however, this difference was not significant, $F(1,120) = 1.56$, $p = 0.21$, $\eta^2 = 0.01$. There was no significant effect of pretest, $F(1,120) = 2.63$, $p = 0.11$ $\eta^2 = 0.02$. Thus, hypothesis H2a is rejected. Similarly, there were no significant correlations between awareness of knowledge gaps after completing the task and posttest scores for the experimental group IA, $r = 0.04$, $p = 0.73$, $n = 62$, and for the experimental group WS, $r = -0.05$, $p = 0.71$, $n = 61$. Therefore, hypothesis H2b cannot be confirmed.

Table 5 presents the partial correlations between variables of solution quality and the posttest for the entire sample and for the two experimental groups. For all partial correlations, the score of the pretest served as the control variable. For the overall sample, there were no significant correlations between the variables of solution quality and posttest scores. There were no significant correlations between the posttest scores and the content-related solution quality variables for the WS experimental group. In contrast, for IA experimental group, the number of categories mentioned correlated

TABLE 3 Selected categories of the facet monitoring of the content-related solution quality.

Feature	Category	TE	Description
			This category means that ...
Surface feature	Communication & behavior of the teacher	TE1	… the teacher often appears annoyed and emphasizes the misbehavior of the students.
		TE2	… the teacher communicates in a friendly yet decisive manner and praises desirable behavior.
Deep feature	Dealing with disruptions	TE1	… the teacher does not manage to maintain a flow through her disruptive interventions.
		TE2	… the teacher is able to maintain a flow through purposeful, undramatic actions.

TE = teaching example. Adapted from Wedde et al. (2022), p. 533.

TABLE 4 Means and standard deviations for the entire sample for pre- and posttest.

Variable	IA	WS	Overall
	M (*SD*)	*M* (*SD*)	*M* (*SD*)
Pretest	0.31 (0.19)	0.37 (0.18)	0.34 (0.18)
Posttest	0.35 (0.21)	0.43 (0.19)	0.39 (0.20)

A coefficient of agreement from 0 to 1 is used for the pretest and posttest. $N = 127$, invention activity (IA): $n = 63$, worked solution (WS): $n = 64$.

with posttest scores negatively and almost significantly with a small effect. The deep features and posttest scores correlated negatively significantly with a small effect. Thus, hypothesis H3a is rejected whereas the hypotheses H3b and H3c can be partially confirmed. The posttest results correlated significantly with the depth of comparison for the experimental group WS (medium effect) but not for the experimental group IA, where there was no significant correlation between the posttest and the depth of comparison. Therefore, hypothesis H3d can only be partially confirmed.

6. Study 2

6.1. Hypotheses

We were able to demonstrate in Study 1 that both learning formats promoted professional vision of classroom management. However, Study 1 did not show which of the two experimental groups was more likely to initiate long-term skill acquisition. Hence, we conducted Study 2 with a new sample. To gain detailed insights into these task formats, we included a follow-up survey to examine the extent to which sustained learning can be fostered by the two task formats. According to the generation effect, it can be assumed that working on an invention activity is more likely to promote sustainable learning than studying a worked solution (Bertsch et al., 2007). However, we assume a high element interactivity for our task material, which is why it is questionable whether the generation effect would occur (Chen et al., 2015). Given the inconsistent evidence from studies on this issue, we also formulated non-directional hypotheses.

> *H1*: The experimental group WS achieves higher scores in the posttest than the experimental group IA after controlling for pretest differences.
>
> *H2*: The experimental groups IA and WS differ in the follow-up after controlling for pretest differences.

6.2. Method

6.2.1. Sample

The total sample for Study 2 consisted of 138 students in their first semester in the teacher-training program for secondary schools at the University of Kassel, Germany. Seventy-one students were randomly assigned to experimental group IA (47.9% female; age: $M = 21.8$, $SD = 4.2$). Sixty-seven students were randomly assigned to the experimental group WS (49.3% female; age: $M = 20.9$, $SD = 2.9$). The lecture took place in presence. All students were enrolled in a teacher training program for secondary schools.

Complete data sets were available from 97 participants for all three measurement time points. As in Study 1, a video filter was used to exclude deviant cases. The filter reduced the sample to 54 students (IA: $n = 24$, WS: $n = 30$). The sample dropout corresponds to 44.3% of the sample, with 26 cases dropping out of the experimental group IA and 17 cases dropping out of the experimental group WS. This difference was not significant, χ^2 (1) = 2.46, $p = 0.12$. The groups did not differ regarding age, $F(1,95) = 0.08$, $p = 0.78$, $\eta^2 = 0.00$, and regarding gender, χ^2 (1) = 3.41, $p = 0.07$. We examined the extent to which the two groups differed with respect to various variables at posttest. The excluded and included groups were found to differ almost significantly in the posttest. No significant differences were found between the two groups regarding the other variables (see Table 6; see Section 6.2.2 for a description of these variables). Separate tests of these differences for the two experimental groups revealed no significant differences for either experimental group regarding the individual variables.

6.2.2. Research design, treatment and instruments

The research design and treatment were identical to those in Study 1. In addition, twelve weeks after completing the posttest, all students were invited to participate in the follow-up. The students completed the follow-up independently. The video-based online test used for the pre- and posttest was also used for the follow-up (see Section 5.2.4). See Figure 2 for an overview of the research design.

For a more detailed dropout analysis, we included additional variables in the Study 2 posttest that had not been collected in Study 1. Situational interest in classroom management was assessed via seven items in the posttest (adapted from Glogger-Frey et al., 2015, e.g., "Having learned about classroom management strategies is *important* to me"). These items were rated by students on a six-point scale ranging from 1 "strongly disagree" to 6 "strongly agree". The scale showed satisfactory reliability ($\alpha = 0.70$). Similarly, we used two scales to assess test motivation in the posttest (adapted from Bürgermeister et al., 2011). Intrinsic motivation was assessed by students via three items (e.g., "I enjoyed working on the teaching analysis"; $\alpha = 0.75$). Whether students were unmotivated was

TABLE 5 Partial correlations between the variables of solution quality and the posttest controlling for the pretest.

Solution quality	Variable	Posttest		
		IA (n = 63)	WS (n = 64)	Overall (N = 127)
Analytical	Depth of comparison (H3d)	−0.14	0.29*	0.10
Content-related	Surface features (H3a)	−0.07	−0.16	−0.07
	Deep features (H3b)	−0.27*	0.03	−0.03
	Number of mentioned categories (H3c)	−0.21⁺	−0.05	−0.05

IA = invention activity; WS = worked solution, $^{+}p < 0.10$, $^{*}p < 0.05$.

TABLE 6 Differences between groups included or excluded by the video filter.

Variable	M (SD)		U	z	p
	Through video filter ...				
	... excluded n = 43	... included n = 54			
Pretest	0.33 (0.16)	0.32 (0.17)	916.50	−1.32	0.19
Posttest	0.31 (0.18)	0.39 (0.19)	849.00	−1.74	0.08
Situational interest t_2	5.03 (0.69)	5.24 (0.50)	1111.50	−0.12	0.91
Amotivated motivation t_2	1.49 (0.52)	1.42 (0.41)	1054.00	−0.48	0.63
Intrinsic motivation t_2	2.60 (0.66)	2.69 (0.58)	1082.00	−0.31	0.76

A coefficient of agreement from 0 to 1 is used for the pretest and posttest. Since the individual variables are not normally distributed (Shapiro-Wilk test, $p < 0.05$), Mann-Whitney-U tests were calculated to determine differences between the groups.

reported by students in four items (e.g., "I thought the teaching analysis questions were stupid"; $\alpha = 0.70$). All items on the test motivation were rated on a scale from 1 "not at all true" to 4 "completely true".

6.2.3. Data analysis

Means and standard deviations for descriptive analysis were calculated for all variables. To determine the differences between experimental groups at posttest and follow-up, ANCOVAs were calculated. The posttest and follow-up served as the dependent variable and the experimental groups as the independent variable. Before ANCOVAs were calculated, the prerequisites were checked. A normal distribution can be assumed for the individual variables for both experimental groups. The variances of the variables were homogeneous (posttest: $p = 0.08$, follow-up: $p = 0.17$). Although the two experimental groups did not differ significantly in the pretest, $F(1,52) = 0.20$, $p = 0.66$, $\eta^2 = 0.004$, we included the pretest as a control variable to increase statistical power. SPSS 27 was used for data analysis.

An *a priori* power analysis had already been calculated for Study 1 and was identical for this study (see Section 5.2.5). The sample size of this study of 54 student teachers has a smaller statistical post-hoc power of $(1\text{-}\beta) = 0.44$.

6.3. Results

Table 7 shows the descriptive statistics of the two experimental groups and the entire sample for pretest, posttest and follow-up. Both experimental groups improved at the posttest compared with the pretest, but the two experimental groups did not differ significantly in the posttest, $F(1,51) = 0.06$ $p = 0.82$, $\eta^2 = 0.00$. There was a highly significant effect of pretest, $F(1,51) = 31.74$, $p < 0.001$, $\eta^2 = 0.38$. Therefore, hypothesis H1 is rejected. In the follow-up, the experimental group IA performed unchanged from the posttest, whereas the WS group's test performance declined and was even below the pretest value. An ANCOVA showed a highly significant effect of experimental group, $F(1,51) = 11.61$, $p = 0.001$, $\eta^2 = 0.19$, and a highly significant effect of pretest, $F(1,51) = 28.87$, $p < 0.001$, $\eta^2 = 0.36$. Thus, hypothesis H2 can be confirmed.

7. Discussion

7.1. Summary and interpretation of results

The objective of these studies was to determine which of two task formats, IA and WS, was preferable for promoting professional vision amongst first-year students. The two formats were compared using contrasting teaching examples. To this end, we determined that both experimental groups improved similarly from pre- to posttest and that there were no significant differences between the two experimental groups at posttest in either Study 1 or Study 2 (RQ1). Initially, this result is, indeed, positive, as both experimental groups demonstrated higher professional vision of classroom management. However, neither of the posttest in our studies showed an advantage for one of the two task formats. This finding contrasts to those of other studies that found WS to be the superior task format, especially for tasks with a high element interactivity, which we assumed for the learning material used in this study (Chen et al., 2015; Glogger-Frey et al., 2015; Roelle and Berthold, 2015; Glogger-Frey et al., 2022).

To investigate how sustainable the skill acquisition was, we conducted a follow-up in Study 2. Here, the results of the experimental group WS did not differ significantly from the results of the experimental group IA at posttest. However, both groups differed highly significant at follow-up; the experimental group IA demonstrated consistent results from posttest to follow-up and the results of the experimental group WS decreased from posttest to follow-up (RQ3). This finding would indicate a generation effect (Bertsch et al., 2007). As a result, for the long-term promotion of professional vision of classroom management, IA is the preferable format. It can be assumed that, since

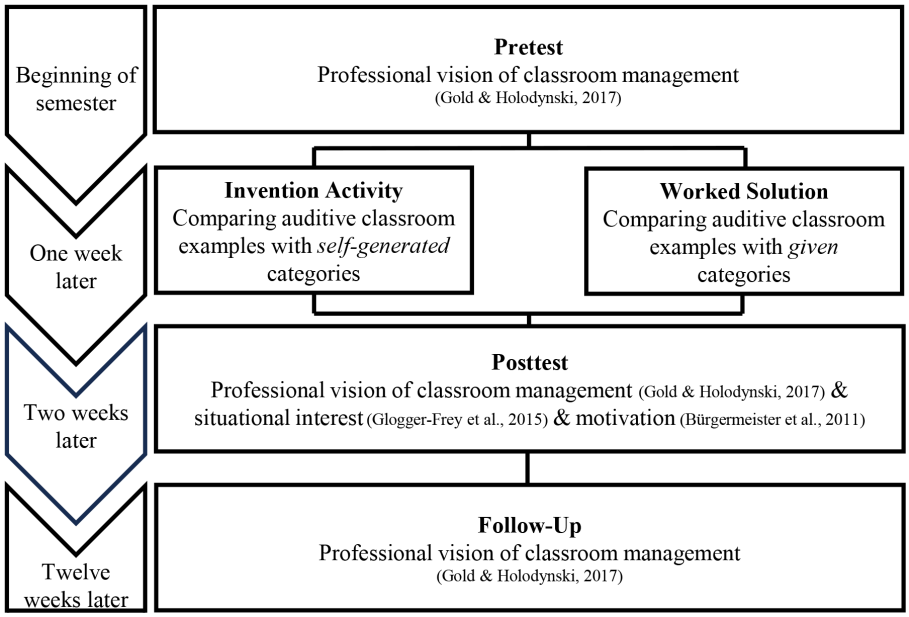

FIGURE 2
Research design of study 2.

TABLE 7 Means and standard deviations for pre- and posttest and follow-up.

Variable	IA	WS	Overall
	M (*SD*)	*M* (*SD*)	*M* (*SD*)
Pretest	0.31 (0.17)	0.33 (0.16)	0.32 (0.17)
Posttest	0.38 (0.17)	0.40 (0.21)	0.39 (0.19)
Follow-up	0.38 (0.17)	0.27 (0.15)	0.32 (0.17)

A coefficient of agreement from 0 to 1 is used for the pretest and posttest. $N = 54$, invention activity (IA): $n = 24$, worked solution (WS): $n = 30$.

the group members developed the categories themselves, deeper cognitive processes were triggered, which resulted in their longer-term ability to perceive classroom management relevant events professionally. The results of this study suggest that when element interactivity is high, more sustained learning occurs with an invention activity.

A previous study found that students in an IA experimental group perceive more knowledge gaps while working on the task than those in a WS experimental group (Glogger-Frey et al., 2015). We could not replicate this result with Study 1, as there were no differences between the two experimental groups. Similarly, awareness of knowledge gaps was not related to the posttest.

The correlations between the variables of the analytical and content-related solution quality of the task showed a mixed picture for the two experimental groups in Study 1 (RQ2). The analytical solution quality of the experimental group WS was positively significantly related to the posttest. This finding may suggest that working through the steps of an effective comparison could initiate professional vision (Wedde et al., under review). Yet, the variables of the content-related solution quality were not significantly related to the posttest of the experimental group WS. Thus, no significant correlations between the number of surface features and the posttest could be demonstrated for the experimental group WS or for the experimental group IA (H3a). Furthermore, no significant correlation existed between analytical solution quality and posttest for experimental group IA (H3d). Nevertheless, two (almost) negatively significant correlations existed between the number of mentioned categories (H3c) as well as the number of deep features (H3b) and the posttest for experimental group IA. Thus, the fewer categories or the fewer deep features used for comparison, the better students performed on the posttest. This finding would suggest a productive failure effect for experimental group IA (Kapur and Bielaczyc, 2012).

With this result, it would have been expected that the experimental group IA would perceive correspondingly more knowledge gaps after task completion, which would have been a possible explanation for the results of this study. The assumption would be that students in experimental group IA are more likely to fail to develop an optimal solution to the task. This failure would lead them to perceive correspondingly more knowledge gaps when working on the task. During instruction, they would then be able to fill previously perceived knowledge gaps, which could lead to more sustained learning as well as networked and organized knowledge. For the experimental group WS, this explanation would mean that, although the students achieved a higher solution quality during the problem-solving phase (Wedde et al., 2022, under review), they perceive fewer knowledge gaps after working on the task and are thus less able to fill these knowledge gaps during instruction, which is why long-term learning is less likely to occur.

7.2. Limitations

The low level of students' awareness of knowledge gaps in Study 1 cannot contribute to the explanation of these results. This failure may also be due to collection through the standardized items used. In the future, less abstract items or interviews could be used to collect data on the awareness of knowledge gaps following the task completion. According to the statistical power analysis, the estimated sample size for a medium effect size had to be larger than the actual sample in Study 2, limiting the

significance of Study 2's findings. Sample dropout is another limitation of the two studies. In Study 1, the group excluded by the video filter demonstrated higher professional vision of classroom management in the pretest. In Study 2, the group excluded by the video filter demonstrated a lower professional vision in the posttest than the included group. This difference raises the question of how the results of the two studies would have been if these sub-samples had compliant video times and thus did not need to be excluded. Furthermore, no control group was used in this study, so it cannot be ruled out that the results are a test–retest effect. Thus, no results are available on how a control group that had not worked on a comparative task would have developed in their professional vision of classroom management. Additionally, the sample was drawn from only one institution, which also limits the study's generalizability.

7.3. Conclusion and implications for further research

Overall, our study was able to show that there is still unclarity in educational science as to which of the two task formats, IA and WS, is preferable. However, this preference also depends significantly on the learning object. When it comes to developing conceptual knowledge on a topic, WS may be superior to IA (Glogger-Frey et al., 2015, 2022). However, if a situation-specific cognitive skill based on declarative and conceptual knowledge is to be fostered beyond conceptual knowledge, IA may well be the preferable learning format, as our studies have shown. To address in detail the question of the appropriate task format, further studies should examine the extent to which comparing contrasting teaching examples specifically promotes conceptual knowledge of classroom management. Likewise, constructs of affect motivation, which are also the basis of professional vision, could also be included in further analyses (Blömeke and Kaiser, 2017). Additionally, the results showed that comparing contrasting auditive teaching examples positively influences student teachers' professional vision of classroom management. Future research could also compare auditive contrasting cases with video-based cases to determine which cases are more likely to initiate student teachers' professional vision at the beginning of their studies with low cognitive load. However, further studies should examine the extent to which the productive failure effect can be replicated. Performing such studies could address the question of how the effect occurred in the context of our study, although previous studies in the field of educational science found contrary results (Glogger-Frey et al., 2015, 2022). In addition, the effect found in Study 2 would have to be verified in a follow-up with a larger sample and with students from different institutions.

In Study 2, the IA experimental group had better results in the follow-up test. The reasons for this difference between the two groups are currently speculative. A further study should examine the differences in more detail. Such a study could, in particular, investigate whether deeper cognitive processes had occurred in the IA group, a suggestion put forward but not substantiated in our previous studies (Wedde et al., 2022, under review). For this purpose, a further study would have to be conducted to determine the reasons for the better performance of the IA experimental group. Nevertheless, we were able to demonstrate that auditive teaching examples are appropriate for promoting professional vision of classroom management among first-year students. These examples reduced the complexity of the analysis during the acquisition phase. Whether this reduction in complexity also succeeds in bringing the significance of language into focus when novices analyze teaching situations must be examined in further studies. To this end, it would be necessary to specifically examine in the posttest and follow-up to what extent the study participants refer to language elements in their professional vision and use these to interpret the visuals (Sabers et al., 1991).

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical approval was not required for the study involving human participants in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was provided by the participants in accordance with the national legislation and the institutional requirements.

Author contributions

AB and DB supervised the project. AB, DB and SW contributed to the conception and design of the study. AB, BG and SW were responsible for the data collection. SW performed the statistical analysis and was responsible for writing the original draft. AB, BG and DB provided critical feedback and edited the manuscript. All authors contributed to manuscript revision, read and approved the submitted version.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

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.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2023.1257681/full#supplementary-material

References

Alfieri, L., Nokes-Malach, T. J., and Schunn, C. D. (2013). Learning through case comparisons: a meta-analytic review. *Educ. Psychol.* 48, 87–113. doi: 10.1080/00461520.2013.775712

Apthorp, H. (2010). "Identifying similarities and differences" in *Classroom instruction that works*. eds. A. D. Beesley and H. S. Apthorp (Denver, CO: Mid-continent Research for Education and Learning (McREL)). S. 14–31.

Ashman, G., Kalyuga, S., and Sweller, J. (2020). Problem-solving or explicit instruction: which should go first when element interactivity is high? *Educ. Psychol. Rev.* 32, 229–247. doi: 10.1007/s10648-019-09500-5

Berliner, D. C. (2001). Learning about and learning from expert teachers. *Int. J. Educ. Res.* 35, 463–482. doi: 10.1016/S0883-0355(02)00004-6

Bertsch, S., Pesta, B. J., Wiscott, R., and McDaniel, M. A. (2007). The generation effect: a meta-analytic review. *Mem. Cogn.* 35, 201–210. doi: 10.3758/BF03193441

Blömeke, S., Gustafsson, J.-E., and Shavelson, R. J. (2015). Beyond dichotomies. Competence viewed as a continuum. *Z. Psychol.* 223, 3–13. doi: 10.1027/2151-2604/a000194

Blömeke, S., and Kaiser, G. (2017). "Understanding the development of teachers' professional competencies as personally, situationally and socially determined" in *International handbook of research on teacher education*. eds. J. D. Clandinin and J. Husu (London: SAGE Publishers), 783–802.

Borko, H., and Livingston, C. (1989). Cognition and improvisation: differences in mathematics instruction by expert and novice teachers. *Am. Educ. Res. J.* 26, 473–498. doi: 10.3102/00028312026004473

Borko, H., and Putnam, R. T. (1996). "Learning to teach," in *Handbook of educational psychology*. eds. D. C. Berliner and R. C. Calfee (New York: Macmillan). 673–708.

Bürgermeister, A., Kampa, M., Rakoczy, K., Harks, B., Besser, M., Klieme, E., et al. (2011). *Dokumentation der Befragungsinstrumente des Laborexperimentes im Projekt conditions and consequences of classroom assessment (Co²CA)*. Frankfurt am Main: DIPF.

Calderhead, J., and Robson, M. (1991). Images of teaching: student teachers' early conceptions of classroom practice. *Teach. Teach. Educ.* 7, 1–8. doi: 10.1016/0742-051X(91)90053-R

Chaplain, R. (2008). Stress and psychological distress among trainee secondary teachers in England. *Educ. Psychol.* 28, 195–209. doi: 10.1080/01443410701491858

Chen, O., Kalyuga, S., and Sweller, J. (2015). The worked example effect, the generation effect, and element interactivity. *J. Educ. Psychol.* 107, 689–704. doi: 10.1037/edu0000018

Chi, M., Dohmen, I., Shemwell, J. T., Chin, D. B., Chase, C. C., and Schwartz, D. L. (2012). Seeing the forest from the trees: A comparison of two instructional models using contrasting cases. *Paper presented at the annual meeting of the American Educational Research Association* (BC, Canada: Vancouver).

Chi, M. T. H., Feltovich, P. J., and Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cogn. Sci.* 5, 121–152. doi: 10.1207/s15516709cog0502_2

Chin, D. B., Chi, M., and Schwartz, D. L. (2016). A comparison of two methods of active learning in physics: inventing a general solution versus compare and contrast. *Instr. Sci.* 44, 177–195. doi: 10.1007/s11251-016-9374-0

Emmer, E. T., Evertson, C. M., Clements, B. S., and Worsham, M. E. (1994). *Classroom management for secondary teachers*. Needham Heights, MA: Allyn and Bacon.

Erickson, F. (2007). "Ways of seeing video: toward a phenomenology of viewing minimally edited footage" in *Video research in the learning sciences*. eds. R. Goldman, R. Pea, B. Barron and S. J. Denny (Mahwah, NJ: Lawrence Erlbaum), 145–155.

Evertson, C. M., and Weinstein, C. S. (2006). "Classroom management as a field of inquiry" in *Handbook of classroom management*. eds. C. M. Evertson and C. S. Weinstein (New York, NY: Routledge), 3–15.

Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). G*power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/BF03193146

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

Glogger-Frey, I., Treier, A.-K., and Renkl, A. (2022). How preparation-for-learning with a worked versus an open inventing problem affect subsequent learning processes in pre-service teachers. *Instr. Sci.* 50, 451–473. doi: 10.1007/s11251-022-09577-6

Gold, B., and Holodynski, M. (2017). Using digital video to measure the professional vision of elementary classroom management: test validation and methodological challenges. *Comput. Educ.* 107, 13–30. doi: 10.1016/j.compedu.2016.12.012

Gold, B., Pfirrmann, C., and Holodynski, M. (2021). Promoting professional vision of classroom management through different analytic perspectives in video-based learning environments. *J. Teach. Educ.* 72, 431–447. doi: 10.1177/0022487120963681

Guest, G., Mac Queen, K. M., and Namey, E. E. (2011). *Applied thematic analysis*. Thousand Oaks, California: SAGE.

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

König, J., Blömeke, S., Klein, P., Suhl, U., Busse, A., and Kaiser, G. (2014). Is teachers' general pedagogical knowledge a premise for noticing and interpreting classroom situations? A video-based assessment approach. *Teach. Teach. Educ.* 38, 76–88. doi: 10.1016/j.tate.2013.11.004

König, J., and Kramer, C. (2016). Teacher professional knowledge and classroom management: on the relation of general pedagogical knowledge (GPK) and classroom management expertise (CME). *ZDM* 48, 139–151. doi: 10.1007/s11858-015-0705-4

König, J., Santagata, R., Scheiner, T., Adleff, A.-K., Yang, X., and Kaiser, G. (2022). Teacher noticing: a systematic literature review of conceptualizations, research designs, and findings on learning to notice. *Educ. Res. Rev.* 36:100453. doi: 10.1016/j.edurev.2022.100453

Korpershoek, H., Harms, T., Boer, H. D., van Kuijk, M., and Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Rev. Educ. Res.* 86, 643–680. doi: 10.3102/0034654315626799

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

Kramer, C., König, J., Strauß, S., and Kaspar, K. (2020). Classroom videos or transcripts? A quasi-experimental study to assess the effects of media-based learning on pre-service teachers' situation specific skills of classroom management. *Int. J. Educ. Res.* 103:101624. doi: 10.1016/j.ijer.2020.101624

Kunter, M., and Voss, T. (2013). "The model of instructional quality: a multicriteria analysis" in *Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project*. eds. M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss and M. Neubrand (Boston, MA: Springer US), 97–124.

Landrum, T. J., and Kauffman, J. M. (2006). "Behavioral approaches to classroom management" in *Handbook of classroom management*. eds. C. M. Evertson and C. S. Weinstein (New York, NY: Routledge), 47–71.

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). Knowing what you don't know makes failure productive. *Learn. Instr.* 34, 74–85. doi: 10.1016/j.learninstruc.2014.08.004

Loibl, K., and Rummel, N. (2014b). 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., Tillema, M., Rummel, N., and van Gog, T. (2020). The effect of contrasting cases during problem solving prior to and after instruction. *Instr. Sci.* 48, 115–136. doi: 10.1007/s11251-020-09504-7

Low, R., and Sweller, J. (2009). "The modality principle in multimedia learning" in *The Cambridge handbook of multimedia learning*. ed. R. E. Mayer (New York, NY: Cambridge University Press), 147–158.

Martin, N. K., and Sass, D. A. (2010). Construct validation of the behavior and instructional management scale. *Teach. Teach. Educ.* 26, 1124–1135. doi: 10.1016/j.tate.2009.12.001

Marzano, R. J., Pickering, D. J., and Pollok, J. E. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement*. Alexandria, VA: Association for Supervision and Curriculum Development.

Mayer, R. E. (2009). "Cognitive theory of multimedia learning" in *The Cambridge handbook of multimedia learning*. ed. R. E. Mayer (New York, NY: Cambridge University Press), 31–48.

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 *Human fallibility*. eds. J. Bauer and C. Harteis (Dordrecht: Springer Netherlands), 53–70.

Plöger, W., Krepf, M., Scholl, D., and Seifert, A. (2020). Analytical competence of teachers: assessing the construct validity by means of mixed methods and drawing consequences for teacher education. *Teach. Educ. Q.* 47, 134–158.

Renkl, A. (2014). Toward an instructionally oriented theory of example-based learning. *Cogn. Sci.* 38, 1–37. doi: 10.1111/cogs.12086

Roelle, J., and Berthold, K. (2015). Effects of comparing contrasting cases on learning from subsequent explanations. *Cogn. Instr.* 33, 199–225. doi: 10.1080/07370008.2015.1063636

Roll, I., Aleven, V., and Koedinger, K. R. (2011). Outcomes and mechanisms of transfer in invention activities. Proceedings of the Annual Meeting of the Cognitive Science Society 33, 2824–2829.

Sabers, D. S., Cushing, K., and Berliner, D. C. (1991). Differences among teachers in a task characterized by simultaneity, multidimensional, and immediacy. *Am. Educ. Res. J.* 28, 63–88. doi: 10.3102/00028312028001063

Schwartz, D. L., Chase, C. C., Oppezzo, M. A., and Chin, D. B. (2011). Practicing versus inventing with contrasting cases: the effects of telling first on learning and transfer. *J. Educ. Psychol.* 103, 759–775. doi: 10.1037/a0025140

Seidel, T., and Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: the role of theory and research design in disentangling meta-analysis results. *Rev. Educ. Res.* 77, 454–499. doi: 10.3102/0034654307310317

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

Sherin, M. G. (2007). "The development of teachers' professional vision in video clubs" in *Video research in the learning sciences*. eds. R. Goldman, R. Pea, B. Barron and S. J. Denny (Mahawah, New Jersey: Lawrence Erlbaum Associates, Publishers), 383–395.

Sherin, M. G., and van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *J. Teach. Educ.* 60, 20–37. doi: 10.1177/0022487108328155

Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., and Sugai, G. (2008). Evidence-based practices in classroom management: considerations for research to practice. *Educ. Treat. Child.* 31, 351–380. doi: 10.1353/etc.0.0007

Sinha, T., Kapur, M., West, R., Catasta, M., Hauswirth, M., and Trninic, D. (2021). Differential benefits of explicit failure-driven and success-driven scaffolding in problem-solving prior to instruction. *J. Educ. Psychol.* 113, 530–555. doi: 10.1037/edu0000483

Stahnke, R., and Blömeke, S. (2021). Novice and expert teachers' situation-specific skills regarding classroom management: what do they perceive, interpret and suggest? *Teach. Teach. Educ.* 98:103243. doi: 10.1016/j.tate.2020.103243

Star, J. R., Lynch, K., and Perova, N. (2011). "Using video to improve preservice mathematics teachers' abilities to attend to classroom features" in *Mathematics teacher noticing: Seeing through teachers' eyes*. eds. M. G. Sherin, V. Jacobs and R. Philipp (New York, NY: Routledge), 117–133.

Stürmer, K., Könings, K. D., and Seidel, T. (2013). Declarative knowledge and professional vision in teacher education: effect of courses in teaching and learning. *Br. J. Educ. Psychol.* 83, 467–483. doi: 10.1111/j.2044-8279.2012.02075.x

Sweller, J., van Merrienboer, J., and Paas, F. (1998). Cognitive architecture and instructional design. *Educ. Psychol. Rev.* 10, 251–296. doi: 10.1023/a:1022193728205

Syring, M., Bohl, T., Kleinknecht, M., Kuntze, S., Rehm, M., and Schneider, J. (2015). Videos oder Texte in der Lehrerbildung? Effekte unterschiedlicher Medien auf die kognitive Belastung und die motivational-emotionalen Prozesse beim Lernen mit Fällen. *Z. Erzieh.* 18, 667–685. doi: 10.1007/s11618-015-0631-9

van den Bogert, N., van Bruggen, J., Kostons, D., and Jochems, W. (2014). First steps into understanding teachers' visual perception of classroom events. *Teach. Teach. Educ.* 37, 208–216. doi: 10.1016/j.tate.2013.09.001

van Tartwijk, J., P. D. Brok, Veldman, I., and Wubbels, T. (2009). Teachers practical knowledge about classroom management in multicultural classrooms. *Teach. Teach. Educ.*, 25:453–460. doi: 10.1016/j.tate.2008.09.005

Wang, M. C., Haertel, G. D., and Walberg, H. J. (1993). Toward a knowledge base for school learning. *Rev. Educ. Res.* 63, 249–294. doi: 10.3102/00346543063003249

Weaver, J. P., Chastain, R. J., DeCaro, D. A., and DeCaro, M. S. (2018). Reverse the routine: problem solving before instruction improves conceptual knowledge in undergraduate physics. *Contemp. Educ. Psychol.* 52, 36–47. doi: 10.1016/j.cedpsych.2017.12.003

Weber, K. E., Gold, B., Prilop, C. N., and Kleinknecht, M. (2018). Promoting pre-service teachers' professional vision of classroom management during practical school training: effects of a structured online- and video-based self-reflection and feedback intervention. *Teach. Teach. Educ.* 76, 39–49. doi: 10.1016/j.tate.2018.08.008

Wedde, S., Busse, A., and Bosse, D. (2021). Zur Wirksamkeit von Invention Actitivities auf das Lernen von Lehramtsstudierenden [The effectiveness of invention activities on student teacher learning]. *Zeitschrift für Hochschulentwicklung.* 16, 35–54. doi: 10.3217/zfhe-16-01/03

Wedde, S., Busse, A., and Bosse, D. (2022). Content-related solution quality in invention activities and worked solutions – promoting the professional vision of classroom management. *Int. J. Mod. Educ. Stud.* 6, 523–546. doi: 10.51383/ijonmes.2022.245

Wedde, S., Busse, A., and Bosse, D. (under review). *Comparing contrasting cases in problem-solving first activities – Insights into analytical solution quality.*

Wiedmann, M., Leach, R. C., Rummel, N., and Wiley, J. (2012). Does group composition affect learning by invention? *Instr. Sci.* 40, 711–730. doi: 10.1007/s11251-012-9204-y

Wolff, C. E., van den Bogert, N., Jarodzka, H., and Boshuizen, H. P. A. (2015). Keeping an eye on learning: differences between expert and novice teachers' representations of classroom management events. *J. Teach. Educ.* 66, 68–85. doi: 10.1177/0022487114549810