



The impact of direct instruction in a problem-based learning setting. Effects of a video-based training program to foster preservice teachers' professional vision of critical incidents in the classroom

Victoria Luise Barth^{a,*}, Valentina Piwovar^a, Irina Rosa Kumschick^b, Diemut Ophardt^c, Felicitas Thiel^a

^aFreie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany

^bTechnische Universität Berlin, Marchstraße 23, 10587 Berlin, Germany

^cPädagogische Hochschule Luzern, Töpferstrasse 10, 6004 Luzern, Switzerland

ARTICLE INFO

Keywords:

Problem-based learning
Classroom management
Preservice teachers
Video
Professional vision
Training program

ABSTRACT

Classroom disruptions are challenging. Problem-based learning (PBL) may help preservice teachers prepare for these situations through self-directed knowledge acquisition or direct instruction. In a first study, we applied a two-group design where students acquired knowledge through either self-directed learning (CG) or direct instruction (EG). Depending on the treatment, we examined differences in knowledge about classroom disruptions and in professional vision (noticing and knowledge-based reasoning). Knowledge was assessed with a multiple-choice test, and professional vision through video case analysis. EG showed higher scores in knowledge than CG and mentioned more knowledge-based reasons. In a second study, pre-post comparison showed increased knowledge and reasoning over time. Noticing did not differ between groups in Study 1 and remained stable in Study 2.

1. Introduction

Preservice teachers struggle to manage student misbehavior and to implement effective strategies for preventing and coping with classroom disruptions when beginning their teaching careers (Aloe, Amo, & Shanahan, 2014; Gallagher, 2009; Melnick & Meister, 2008). A key precondition for preventing and intervening in disruptions is that teachers are able to identify relevant characteristics of the situation (noticing) and draw conclusions regarding their own actions (knowledge-based reasoning) in highly complex classroom situations (Doyle, 1986). If a teacher notices, for instance, that students are engaged in activities unrelated to the lesson, they can intervene quickly to recapture students' attention and get them involved in the lesson again. The skills of noticing and knowledge-based reasoning are subsumed under the term "professional vision" (Sherin, 2001).

Findings on professional vision (Lachner, Jarodzka, & Nückles, 2016) show: Novices tend to notice misbehavior too late or not at all and to misinterpret classroom events (van den Bogert, van Bruggen, Kostons, & Jochems, 2014; Housner & Griffey, 1985). Yet, preservice teachers' professional vision of critical incidents in the classroom is often neglected or not addressed sufficiently in teacher education programs (Christofferson & Sullivan, 2015; Freeman, Simonsen, Briere, & MacSuga-Gage, 2014; van Tartwijk &

* Corresponding author at: Department of Education and Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany.
E-mail address: victoria.barth@fu-berlin.de (V.L. Barth).

Hammerness, 2011; Gold, Förster, & Holodynski, 2013). The challenges inherent in teaching classroom management to preservice teachers exacerbate this problem further (Brophy, 1988).

Practice-based learning environments can foster teachers' professional vision of critical incidents in classroom through case-based learning from video as well as problem-based learning. Case-based learning with videotaped lessons has proven to be a valuable method to enhance both preservice and inservice teachers' classroom management knowledge and skills (Brophy, 1988; Brophy, 2004), including the ability to give clear instructions (Ophardt & Thiel, 2013; Piwowar, Thiel, & Ophardt, 2013) and the professional vision of classroom management (Gold et al., 2013). Problem-based learning (PBL), on the other hand, is an instructional approach that allows preservice teachers to systematically analyze video cases as part of their learning process. While research on PBL in teacher education is sparse, higher education research on PBL has demonstrated its positive impact on relevant outcomes, including its potential to increase motivation and professional skills (Dochy, Segers, van den Bossche, & Gijbels, 2003; Hmelo-Silver, 2004).

Although promising, learning settings that merge PBL and video-based learning are still understudied in the field of teacher education. Recent findings from empirical research (Alfieri, Brooks, Aldrich, & Tenenbaum, 2011; Dochy et al., 2003; Wijnia, Loyens, van Gog, Derous, & Schmidt, 2014) as well as theoretical contributions to the debate on the general usefulness of PBL (Kirschner, Sweller, & Clark, 2006; Sweller, Kirschner, & Clark, 2007) suggest that the traditional PBL approach may not be suitable for novice learners who possess only partial and relatively isolated knowledge in the target domain. For preservice teachers, self-directed and unassisted acquisition, integration, and application of a knowledge base in a PBL cycle may be too cognitively demanding and inexpedient.

The purpose of the current study was twofold. First, we intended to implement and evaluate the effectiveness of a training program to foster preservice teachers' professional vision of critical incidents in the classroom based on PBL. We were particularly interested in whether preservice teachers' training enables them to build and develop their knowledge base on classroom management (especially with regard to preventing and dealing with student misbehavior) and their professional vision of critical incidents in classroom. In order to deepen our understanding of target-group-specific learning settings, we also studied the suitability of a single element of conventional PBL for preservice teachers in an experimental setting in more detail: the self-directed acquisition of new information.

2. Fostering preservice teachers' professional vision of critical incidents in classroom

Dealing with classroom disruptions is a key component of classroom management, a competency that can be understood as comprising "all of a teacher's activities that are aimed at establishing and maintaining social order within the social system of the school class" (Ophardt & Thiel, 2013, 46). Building on the work of Doyle (1984), Ophardt and Thiel (2013) describe how these activities are outlined in a lesson plan, which covers learning goals and methods and a program of action that has to be adapted situationally to the concrete teaching context. A number of strategies can be used to prevent and intervene in disruptions (see Ophardt & Thiel, 2013; Thiel, 2016). First, teachers have to establish an interactional order. This means introducing rules and procedures, encouraging and modifying behavior to conform with the rules, and establishing a working relationship. Within this interactional order, teachers establish a clear program of action (e.g., introducing an assignment) and use verbal, paraverbal, and nonverbal signals to control students' attention and behavior. An important precondition for effective control is that teachers continually have an overview of the entire class and that they demonstrate this to their students (see Kounin, 1970). Guiding lesson flow effectively means maintaining continuous group activation, which includes (Kounin, 1970): keeping all students activated (group alerting), keeping all non-active students on standby (high participation formats), and conveying the impression that learning outcomes could be monitored at any time (encouraging accountability). For effective intervention, it is fundamental that the teacher is able to correctly assess the reason for the disruption, the scope of the disruption, and the potential dynamics of the disruption (see Thiel, 2016). Here, it should be kept in mind that, on the one hand, disruptive behavior of the student has to be stopped (individual focus), and at the same time, the program of action for the class must be maintained (class focus). To effectively stop the disruptive behavior, it is crucial that the teacher selects the appropriate strategy. In the case of a minor rule violation (e.g., whispering to a neighbor) it may make sense not to address it directly but to demonstrate awareness of the incident through eye contact. At the same time, the program of action has to be strengthened by intensifying the attentional control of the entire learning group. As soon as a threat to the program of action arises, the teacher has to intervene, or the disruption may create ripple effects. A minimal intervention (e.g., a brief reminder of the rules) can counter this. Teachers' professional vision is a key component of a successful approach to disruptions (see Wettstein, 2013).

The multidimensionality of the classroom, with different events occurring simultaneously, and the immediate need for classroom management are particularly challenging for teachers (Doyle, 2006). A majority of classroom management issues are also relatively ill-structured, with no single best way to handle them effectively. In a study by LePage et al. (2005), an expert-novice comparison showed that the relevant skills and knowledge to manage classrooms develops predominantly during practical experience. As compared to expert teachers, less experienced teachers have trouble balancing care and control (Nie & Lau, 2009; Weinstein, 1998); they tend to lose the group focus of the whole class when reacting to misbehavior (Thiel, Richter, & Ophardt, 2012; Westerman, 1991) and their attention is more scattered across the classroom (Wolff, Jarodzka, van den Bogert, & Boshuizen, 2016). These differences are ascribed to differences in cognitive structures: Experts have more case-based, better connected, and thus more flexible knowledge (which is organized in schemata) than novices. This allows them to notice the critical incidents in the classroom and choose from a broad variety of possible strategies, adapting sensitively to the context at hand (Berliner, 2001; Carter, Cushing, Sabers, Stein, & Berliner, 1988; Spiro, Feltovich, Jacobson, & Coulson, 1991).

Therefore, it seems crucial to initially create preconditions that foster preservice teachers' professional vision of critical incidents

in classroom. This may pertain to 1) building schemata that enable teachers to identify relevant events in the classroom and 2) enhancing the ability to use this knowledge to interpret these events correctly, to anticipate the further course of action, and to develop suitable action strategies. Professional vision (Sherin & van Es, 2009; Sherin, 2001) supports the two processes of noticing and knowledge-based reasoning: Noticing describes whether teachers pay attention to important events and ignore those that are irrelevant to maintaining the activity flow and guaranteeing learning (e.g., small disturbances that do not jeopardize the course of action and may not cause ripple effects). Knowledge-based reasoning refers to the ways in which teachers on the one hand reason about what they notice and on the other hand develop appropriate strategies to handle situations, based on their knowledge and understanding (Sherin & van Es, 2009, e.g., a teacher praises the effort of a student who failed in order to reinforce motivation), indicating differentiated, integrated, and flexible knowledge (Seidel & Stürmer, 2014). In fact, recent research has demonstrated that working with videos is an effective means of teaching professional vision (Gold et al., 2013; Stürmer, Seidel, & Holzberger, 2016; van den Bogert et al., 2014).

3. Problem-based learning in teacher education

PBL is an umbrella term that encompasses many different instructional approaches and applications. A key underlying principle of PBL is that it anchors learning and teaching in concrete, realistic problems (Hmelo-Silver, 2004; Schmidt, 1994). More specifically, Barrows (1996, p. 5f) identified six core characteristics of PBL learning environments: 1) learning is student-centered 2) learning occurs in small student groups, 3) teachers are facilitators or guides, 4) problems form the organizing focus and stimulus for learning 5) problems are a vehicle for the development of problem-solving skills, and 6) new information is acquired through self-directed learning.

According to these characteristics, PBL is supposed to meet a broad spectrum of instructional goals such as promoting the construction of extensive and flexible knowledge, fostering effective skills in problem solving and collaboration, and influencing intrinsic interest in the subject matter (Barrows, 1996; Hmelo-Silver, 2004; Norman & Schmidt, 1992). A number of meta-analyses indeed identified positive effects favoring PBL when compared with traditional instruction: PBL leads to better program evaluation, to superior skills and clinical performance, and to increased retention of knowledge. However, PBL was found to be equally or even less effective than traditional instruction with regard to academic achievement and acquired knowledge (Albanese & Mitchell, 1993; Dochy et al., 2003; Hmelo-Silver, 2004; Norman & Schmidt, 1992; Vernon & Blake, 1993).

Of particular interest to the present study is Dochy et al.'s (2003) meta-analysis, which was the first study that extended the literature base to other domains than medical education and also constrained its selection of studies to solely (quasi-)experimental designs. In line with previous evidence, they found a small but significant effect on student knowledge favoring traditional instruction. Yet, subsequent moderator analyses revealed that the student expertise level was associated with the variation in effect sizes: Traditional instruction led to a more structured knowledge base in studies with first- and second-year students than PBL, but led to a less structured knowledge base for third-year and advanced students. These findings are in line with a theoretical debate initiated by Kirschner et al. (2006). Based on the “knowledge of human cognitive architecture, expert-novice differences, and cognitive load theory” (p. 75), they argue that PBL is not suitable for the knowledge acquisition of novices—even though it may suit advanced learners or experts, who have already stored relevant information in their long-term memory, which in turn frees up capacity in the working memory that is needed for the problem-solving process. In contrast, learners who lack the “extensive experience” that experts draw on to “quickly select and apply the best procedures for solving problems” (p. 76) are overwhelmed when meeting the high demands of unguided PBL.

Consequently, the construction of a PBL environment involves a careful analysis of both the learning goals (addressed outcomes) and the target group (addressed learners). In order to tailor the PBL model to the developmental level of preservice teachers, additional support in the tradition of direct instruction may be implemented, such as scaffolding, elicited explanations, process worksheets, and worked examples (Alferi et al., 2011; Kirschner et al., 2006; Wijnia et al., 2014). In addition, it seems promising to supply preservice teachers with a “framework for reflecting on what was observed [and] concrete images of innovative and alternative teaching strategies” (Santagata, Zannoni, & Stigler, 2007, p. 125).

4. A newly developed training program to foster preservice teachers' professional vision of critical incidents in the classroom

Based on the current state of the research and previous experience with professional learning about classroom management among inservice teachers (Ophardt & Thiel, 2013; Piwowar et al., 2013), we developed a PBL training program for preservice teachers starting with an introduction to classroom management taught by an instructor. The training program consisted of four sessions in total that lasted 90 min each. At the core of the treatment were two ill-structured (video-based) problem cases for the participants to solve collaboratively in groups of four to five students. The problem cases exemplified severely disrupted classes, unfavorable classroom situations, and poor classroom management in a video sequence: the observed teacher applied common, but rather ineffective, strategies to both prevent and handle misbehavior, maintain a conducive teacher-student relationship, and proactively steer the activity flow; the situations escalated (i.e., a student left the classroom in case 1, a student refused to participate in class in case 2). For each problem case, we also had a corresponding best practice example that served as a case exemplar at the end of the treatment (Moreno & Valdez, 2007; Wijnia et al., 2014). In order to be able to provide matched cases of the same initial classroom scenario (problem case vs. case exemplar), we used scripted video vignettes (Piwowar, Barth, Ophardt, & Thiel, 2017). In a previous study (Piwowar et al., 2017) these scripted video vignettes were rated by preservice teachers ($N = 81$). More than 90%

valued the videos both as a good basis for discussion and reflection and as credible and authentic. Eventually, because we had to assume that participants were unfamiliar with PBL, the problem-solving process was structured to follow a process worksheet (Alfieri et al., 2011).

The training program started with two lecture-like sessions that provided a systematic framework on heuristics, including theoretical terms and schemata on questions of classroom management and dealing with disruptions. In the third session, the students formed small groups and watched the two problem (“dysfunctional”) video cases. For each case, they proceeded with a three-step worksheet, which included 1) recording relevant events 2) evaluating the teacher behavior with reference to theoretical and empirical research, and 3) developing alternative, favorable teacher strategies. In the fourth and final session, participants had the opportunity to compare their identified alternative teacher behaviors to those in the respective case exemplars (“functional” cases) and discuss the acquired strategies.

5. Objectives

Since PBL has been shown to be a useful approach in higher education and, more specifically, teacher preparation (e.g., Choi & Yang, 2011; Edwards & Hammer, 2006; Golightly & Raath, 2015; Hushman & Napper-Owen, 2011; Yoon, Woo, Treagust, & Chandrasegaran, 2014; Zhang, Lundeberg, Koehler, & Eberhardt, 2011), we developed the described PBL environment using video cases in a secondary teacher preparation course on the topic of classroom disruptions. PBL also has elements of direct instruction, so although we were convinced of its usefulness in preparing teachers for the classroom, we assumed that it should not be implemented without specific adaptations in this phase of teacher training. Novice teachers, who are not familiar with the PBL cycle, may experience difficulties in acquiring new information through self-directed learning and may therefore be unsuccessful in handling their problem case (Wijnia et al., 2014). Thus, the overarching goal was to gain a broad understanding of the usefulness of this newly developed PBL environment, which contained two research questions. First: Is a training program that merges direct instruction and PBL effective with regard to preservice teachers’ professional vision of critical incidents in the classroom? And second: Is this training program superior to a more traditional PBL treatment starting with a self-study phase?

Therefore, we conducted two experimental studies (see Fig. 1). Study one served to explicitly examine the second question and compared two treatment conditions: Participants in the control group (CG) acquired knowledge through self-study in line with Barrows’ (1996) core characteristics of PBL settings. In the experimental group (EG), on the other hand, relevant theories, strategies, and principles of classroom management were systematically introduced by a university teacher using direct instruction (contradicting Barrows’ sixth core characteristic). While the results of a similar study published in Kumschick et al. (2017) revealed higher content knowledge, higher identified motivation, and higher teacher self-efficacy in classroom management in the EG as opposed to the CG, the present study sought to further complement those results with an in-depth, quantified *qualitative* analysis of aspects of professional vision regarding critical incidents in the classroom. We expected that participants in the EG would outperform

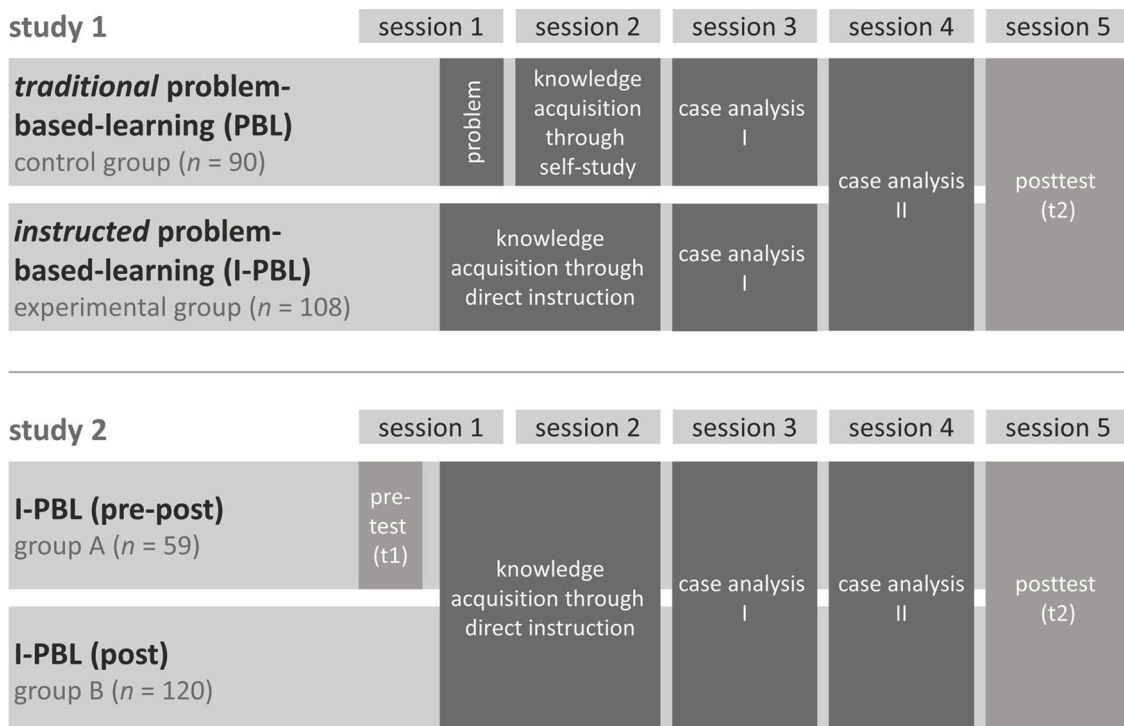


Fig. 1. Research design.

participants in the CG not only with regard to knowledge but also with regard to noticing and reasoning (group difference hypothesis).

In Study 2, we sought to further study *changes* occurring after training participation through pre-post comparison. Based on the results of Study 1, we only conducted the new treatment in the EG described above. However, we used an experimental design in which group A filled out the questionnaires to assess knowledge and professional vision pre-post, and group B filled out the questionnaires only after intervention. This research design allowed us to study enhancement while controlling for potential pretesting effects (affecting the way the treatment works) and retest effects (participants remember their answers and improve performance independently of participation in a treatment, Kulik, Kulik, & Bangert, 1984). We expected that both knowledge and skills (that is, noticing and reasoning) would increase over time (change hypothesis).

6. Study 1

6.1. Method

6.1.1. Sample and study design

The two classroom management trainings were implemented in a university course for $N = 237$ preservice secondary teachers in their first year of master's-level study. Eight study groups were randomly assigned to either the experimental group (EG) or the control group (CG). For this analysis, we only analyzed the data from participants who were part of all measures in post testing with $N = 198$ ($n_{EG} = 108$ and $n_{CG} = 90$). Nine participants who dropped out of the original sample did not take part in video analysis (participation in testing was not mandatory). The EG and the CG were comparable with regard to their sex (58% female) and their age ($M = 26.03$ years, $SD = 4.01$).

6.1.2. Treatments

Whereas the EG took part in the PBL-training described above, the treatment of the CG was rather traditionally problem-based, as the participants gained the relevant knowledge in a more self-directed way. In the first session of the CG, the student groups first watched the problem cases and were then required to record relevant incidents of the video cases collaboratively (step 1 of the process worksheet). Based on this, the groups identified learning topics for the subsequent problem-solving process. In the second session, they learned independently about key theoretical concepts and empirical findings on strategies to prevent and deal effectively with disruptions by reading relevant literature provided in a mandatory reader (contents of the literature were identical to the contents of the systematic introduction in the EG). In the third session, group members first shared and discussed the literature and then proceeded according to the process worksheet (steps 1 and 2). The fourth session was identical to the session in the EG.

Three university teachers were randomly assigned to the seminars; each attended the same number of EG and CG. Thus, teachers were not blind to treatment conditions. The role of the teacher was to facilitate the learning process during problem solving and give the introductory sessions in the EG.

6.1.3. Measures and data analyses

6.1.3.1. Multiple-choice test on classroom management knowledge (MC-test). We conducted a self-developed multiple-choice test with 20 closed-ended items that covered factual knowledge, understanding of linking principles, and an application of both aspects to written practical examples. Participants had to choose one correct out of six possible answers, for example: “Which teacher behavior is most appropriate for handling minor disruptions? Pick a student to repeat what has been said; Engage other students to regulate misbehavior; Reprimand in a timely way; Use gestures and make eye contact (correct); Refer to rules.” The total score resulted from adding up the number of correct answers. Cronbach's alpha was satisfactory at the lower bound with $\alpha = .60$. The development of the instrument was based on a validated observer rating and a student rating for secondary teachers to assess classroom management competencies (Piwovar, 2013). The rating instruments comprise nine theoretical facets of classroom management strategies whose factor structure was tested with exploratory factor analysis (Piwovar, 2013). For test economic reasons, each theoretically modelled facet was represented with about two multiple-choice items. No re-examination of the factor structure was carried out. The development of the MC test contained several pilot phases. First, the items developed were discussed intensively within a group of five experts in this domain. Second, think aloud analyses were applied with two preservice teachers to test item difficulty, the adequacy of phrasing, and the connotations of the distractors. Third, a quantitative indication of congruent test validity was deduced from applying the MC test in a Bachelor seminar that also focused on issues of classroom management ($N = 74$). Since previous studies showed that an increase in knowledge of class management is accompanied by an increase in both self-reported teacher self-efficacy in classroom management (Pfitzner-Eden, Thiel, & Horsley, 2014) and self-reported knowledge (Böttcher-Oschmann, Groß-Ophoff, & Thiel, 2018), we also studied these correlations. We used Pearson's Product Moment Correlation Coefficient to show that the MC test score was correlated with teacher self-efficacy in classroom management ($r = .19$) and two subjective change rating items (“In this seminar I have learned a lot about effective classroom management”, $r = .33$; “In this seminar I did not learn anything really new about classroom management”, $r = -0.43$).

6.1.3.2. Video-based case analysis on professional vision. Teachers have to respond quickly to spontaneous (disruptive) events by recognizing (noticing) relevant situations and evaluating them (reasoning) under pressure to act. To measure the competencies of professional vision in as realistic a way as possible, we took three conditions into account in our measurement method: 1. To visualize the complexity of the teaching situation, we used a video-based case; 2. To simulate the pressure to act, we gave participants

a limited period of time to complete the assignment; 3. Since events also occur just once in situ, they would only view the video case one time. The dependent variables noticing and reasoning were assessed through video-based case analysis with an open-ended question (Barth, 2017). The video prompt was an eight-minute scripted video about coping with classroom disruptions (Piwowar et al., 2017). It was similar to the problem cases used in the training program but displayed a different class and a different teacher. The participants had 40 min to both watch the video clip and then describe what they noticed and how they reasoned about the classroom situation: “In the video you have seen some classroom disruptions. Describe all important aspects of the lesson which have led to the unfavorable course of action. Please justify your choice (if possible, in a theoretically well-founded way) and give for each aspect an example from the teaching situation shown.”

We used qualitative structuring content analysis (Mayring, 2014) to analyze the data. In a first step, the two assessment dimensions noticing and reasoning were defined:

- Noticing: Descriptions of behavior of the teacher or the students
- Reasoning: Evaluations of observed incidents

With the help of four experts in classroom management, we then deductively developed categories to score the participants answers, which were 184 different behavioral characteristics for noticing (e.g., teacher overlooks that a student raised his hand), and 63 theoretical references which are relevant to reasoning about classroom disruptions (e.g. “the monitoring of the teacher is insufficient”, coded as *withitness*, Kounin, 1970). The number of correctly described behavioral characteristics formed the score in noticing, and the number of correctly described theoretical concepts formed the score in reasoning. Eventually, we developed a coding manual containing a definition, anchor samples, and coding rules for each category (Mayring, 2014, p. 95). In a three-day workshop, eight coders (preservice teachers who had attended the seminar previously and earned high grades in the final exams) were trained intensively to apply the coding manual to all given answers. To calculate inter-coder reliability (Cohen’s kappa κ), 10% of the sample ($n = 20$) was double-coded. A mean Cohen’s kappa κ of 0.72 indicated satisfactory inter-coder-reliability (vgl. Fleiss & Cohen, 1973).

As was to be expected, professional vision and knowledge revealed medium-significant correlations ($r = .34$ and $r = .44$ for noticing and reasoning, respectively, see Table 1), the significant correlation of noticing and reasoning was strong ($r = .65$). Age and sex were not meaningfully related to the assessed outcomes ($-0.18 > r < 0.15$).

6.2. Results

Item difficulties of the knowledge MC test were balanced with a minimum of 30% and a maximum of 90% correct answers ($Md = 57\%$, $M = 61\%$). In the video analyses, participants wrote down about 11 correct incidents they had noticed, varying between zero and 27, and listed about five correct different theoretical reasons, varying between zero and 19 (Table 2).

We applied Multivariate Analysis of Variance (MANOVA) with the SPSS24 software to test group differences; the alpha level was set at $p < .05$. To test between-subjects effects, we used the Bonferroni method; the level of significance was corrected to account for the multiple tests (alpha level was set at $p < .02$). Rejection or confirmation of the hypotheses was oriented towards the critical alpha level (that is, statistical significance). However, we also reported effect size partial eta square η_p^2 as an index for the practical relevance of the effects (Richardson, 2011), which is judged to be small at .01, medium at .06, and large at .14 (Cohen, 1988).

Using Pillai’s trace, there was a significant effect of the treatment on knowledge, noticing, and reasoning, $V = 0.053$, $F(3, 190) = 3.542$, $p = .016$, partial $\eta^2 = .053$.

As was hypothesized, the EG scored higher on knowledge, $F(1,192) = 8.453$, $p = .004$, $\eta^2 = .04$, and reasoning, $F(1,192) = 6.581$, $p = .011$, $\eta^2 = .01$, see Table 2. Both effects were small and statistically significant. A small, but non-significant group difference related to noticing, $F(1,192) = 3.896$, $p = .050$, $\eta^2 = .02$.

Table 1
Correlations Between Dependent Variables and Controlling Variables of Study 1 and Study 2.

		1	2	3	4	5	6	7	8	Raven
1	Age		-.18*	.10	-.12	-.14	-.06	-.15*	.00	-.12
2	Sex	-.18*		.03	.04	.01	.15	.14	.06	-.19*
3	Knowledge T1	-	-		.02	.00	.54**	.12	.21	.24
4	Noticing T1	-	-	-		.47**	.09	.58**	.16	-.09
5	Reasoning T1	-	-	-	-		.22	.33*	.26*	.06
6	Knowledge T2	-.02	.15	-	-	-		.27**	.35**	.24**
7	Noticing T2	-.04	.08	-	-	-	.34**		.56**	.19*
8	Reasoning T2	-.09	.02	-	-	-	.44**	.65**		.23**

Note. Correlations of Study 1 are listed below the diagonal, correlations of Study 2 are listed above the diagonal. Sex is coded as 0–male and 1–female, T1–pre-testing, T2–post-testing, $n_{Study1} = 198$, $n_{Study2} = 179$ for post-measures and 59 for pre-measures, * $p < .05$, ** $p < .01$.

Table 2

Post-test Descriptives and One-Way Multivariate Analysis of Variance Summary Table for the Effects of Treatment Group on Knowledge, Noticing and Reasoning in Study One.

		<i>M</i>	<i>SD</i>	MANOVA		
				<i>F</i>	<i>p</i>	η_p^2
Knowledge	EG	64.25	14.80	8.453	.004	.04
	CG	57.95	15.25			
Noticing	EG	11.89	4.78	3.896	.050	.02
	CG	10.51	4.89			
Reasoning	EG	5.85	4.40	6.581	.011	.03
	CG	4.41	3.18			

Note. EG – experimental group with instructed knowledge acquisition, CG – control group with self-study phase for knowledge acquisition. Knowledge – number of correct answers in MC-test in %, Noticing – number of mentioned events, reasoning – number of mentioned theory-based reasons.

7. Study 2

7.1. Method

7.1.1. Sample and study design

In Study 2, the treatment was identical for the whole sample and was like that described for the EG in Study 1. Yet, we ran two evaluation conditions. Because competence testing may lead to pretest effects and thus limit internal validity, we decided to experimentally vary the survey design: Group A filled out questionnaires before and after intervention (pre-post comparison). Group B served as a control group and filled out the questionnaires only in post-testing.

Again, participants were preservice secondary teachers in their first year of a master's program ($N = 179$), with eight study groups that were randomly assigned to either group A ($n_A = 59$) or group B ($n_B = 120$). There was one university teacher for all groups, so the eight study groups took place at different times on different days. After the students were assigned to the individual groups, some students had problems planning their semester courses, so some had to switch to a course on another day. Some study groups therefore had much larger numbers of attendees (e.g., Wednesday courses) than others (e.g., Friday courses), leading to the unequal sample sizes of group A and group B. The groups were comparable with regard to their sex (59% female) and cognitive abilities (assessed by a short form of the raven advanced progressive matrices test by Arthur and Day 1994, with 8.52 out of 12 possible answers). However, the medium age was higher in group A than in group B ($M_A = 28.48$, $SD = 6.59$, $M_B = 26.01$, $SD = 4.13$, $F[1,169] = 9.006$, $p = .003$).

7.1.2. Measures and data analyses

7.1.2.1. Revised multiple-choice test on classroom management knowledge (MC-test-R). We again used the MC test described above, but in a revised version. First, based on an item analysis regarding the item difficulty, we selected 18 (partly modified) items (instead of 20 items previously). Second, we examined the distractors that worked best and reduced the possible answers (from six previously) to five. Cronbach's alpha was satisfactory at $\alpha_{T1} = .60$ and $\alpha_{T2} = .66$.

7.1.2.2. Video-based case analysis on professional vision. The overall procedure to assess noticing and reasoning was identical to the procedure in Study 1. A mean Cohen's kappa κ of 0.74 ($n = 18$) indicated satisfactory inter-coder reliability.

Correlations of knowledge and professional vision were close to zero in pre-testing and approximately medium in post-testing ($r = .27$ and $r = .35$ for noticing and reasoning, respectively, see Table 1). Noticing and reasoning revealed strong correlations in both ($r_{T1} = .47$, $r_{T2} = .56$). We collected further data in post-testing to check divergent validity of the three dependent variables with cognitive abilities (Arthur & Day, 1994) and found small correlations that were about even for knowledge ($r = 0.24$) and reasoning ($r = .23$) and slightly lower for noticing ($r = .19$), as was to be expected. Again, the outcomes were not meaningfully related to the participants' sex or age ($-0.18 > r < 0.15$).

7.2. Results

In a first step, we used data from the experimental design to test whether pre-testing (group A) influenced performance in post-testing by comparison to group B. ANOVAs revealed no group differences in post-testing for knowledge (group A: $M = 63.94\%$, $SD = 17.27$, group B: $M = 61.39\%$, $SD = 15.29$; $F[1,177] = 1.006$, $p = .317$, see Table 3), noticing (group A: $M = 11.44$, $SD = 4.18$, group B: $M = 11.38$, $SD = 4.35$; $F[1,177] = 0.009$, $p = .924$), or reasoning (group A: $M = 4.81$, $SD = 3.15$, group B: $M = 4.43$, $SD = 3.28$; $F[1,177] = .546$, $p = .461$). These results were similar and still non-significant when controlling for age, which differed between the two groups. Consequently, pretesting did not significantly affect scoring in post-testing, and pre-post comparison may be a valid indicator for potential improvement in the studied variables. Thus, we used the data from group A for subsequent change analyses.

Table 3

Pre-Post Descriptives and Repeated Measures Analysis of Variance Summary Table for the Effects on Knowledge, Noticing, and Reasoning in Study 2.

		M	SD	ANOVA		
				F	p	η_p^2
Knowledge	T1	48.21	16.42	56.456	< .001	.49
	T2	63.49	17.27			
Noticing	T1	11.24	5.55	0.114	.737	.00
	T2	11.44	4.18			
Reasoning	T1	3.08	2.14	15.989	< .001	.22
	T2	4.81	3.15			

Note. Knowledge – number of correct answers in MC-test-R in %, Noticing – number of mentioned events, reasoning – number of mentioned theory-based reasons.

In both pre- and post-testing, item difficulties of the multiple-choice knowledge test were balanced (percent of correct answers T1: $Min = 14\%$, $Max = 92\%$, $Md = 50\%$; T2: $Min = 32\%$, $Max = 93\%$, $Md = 70\%$) and total scores varied (T1: $Min = 15\%$, $Max = 80\%$; $M = 51\%$; T2: $Min = 6\%$, $Max = 100\%$, $M = 72\%$). In the video analyses, participants noticed between four and 35 aspects in pretesting ($M = 11.24$) and between three and 22 aspects in post-testing ($M = 11.44$). In average, participants named about three different theoretical reasons at pre-testing ($Min = 0$, $Max = 10$) and about five in post-testing ($Min = 0$, $Max = 13$).

As shown in Table 3, repeated measures ANOVAs revealed significant increases in knowledge ($M1 = 48.21\%$, $SD = 16.42\%$, $M2 = 63.49\%$, $SD = 17.27\%$, $F[1,59] = 56.456$, $p < .001$, $\eta_p^2 = .489$) and reasoning ($M1 = 3.08$, $SD = 2.14$, $M2 = 4.81$, $SD = 3.15$, $F[1,58] = 15.989$, $p = < .001$, $\eta_p^2 = .216$) with large effects (partial eta square η_p^2 served as effect size, Richardson, 2011, which is judged to be small at .01, medium at .06 and large at .14, Cohen, 1988). The ability for noticing did not develop over time ($M1 = 11.24$, $SD = 5.55$, $M2 = 11.44$, $SD = 4.18$, $F[1,58] = .114$, $p = .737$, $\eta_p^2 = .002$).

8. Discussion

Professional vision is a prerequisite for successful teaching (Sherin, 2001) that can be addressed in training at the university level (Gold et al., 2013). But which learning settings are best for preservice teachers with little classroom experience? The present study sought to both implement a training program to foster preservice teachers' professional vision of critical incidents in classroom and to further study the impact of self-directed learning vs. direct instruction of the relevant material in a problem-based learning environment on three cognitive outcomes: knowledge on classroom management, noticing critical incidents in the classroom, and knowledge-based reasoning.

Indeed, the preservice teachers who participated in the newly developed training program enhanced their knowledge and also improved their ability to apply this knowledge in a simulated teaching situation (the video) through knowledge-based reasoning, as could be shown in Study 2. In addition, comparing the two different conditions of the PBL setting in Study 1 revealed that the preservice teachers benefited from our target-group-specific adaptation of the self-study phase: The experimental group, which received a systematic introduction to the relevant knowledge base, showed statistically significantly higher declarative knowledge and knowledge-based reasoning than the control group, which acquired the knowledge in a self-directed way. Thus, we were able to achieve our goal of offering preservice teachers a PBL training program in classroom management that facilitated declarative knowledge, theory-based reasoning, and—as demonstrated in Kumschick et al. (2017)—teacher self-efficacy in a motivating learning environment. Given the outlined challenges to offering suitable learning opportunities, the results reported here are particularly encouraging: we identified small to large effects even though the training consisted of just four sessions. In contrast, the majority of teacher education programs in classroom management and related skills that lead to notable improvements are a full semester in length (Gold et al., 2013; Stough & Montague, 2015). With regard to this economic aspect, and because professional vision is a prerequisite for successful teaching (Sherin, 2001), we implemented the training in seminars that prepare preservice teachers for their first school internship. As classroom disruptions are a special case, it is also conceivable that one could conduct a similar video-based training program with a focus on nonverbal communication to promote classroom management skills.

The training program developed here did not measurably affect the ability of noticing, as was hypothesized: The two training groups did not differ (Study 1), and noticing did not improve over time (Study 2). One possible reason for this finding may be that the training itself did not deal specifically with noticing. While the process worksheet that guided the problem-solving cycle contained the prompt to record noticed incidents in groups, the results of this step were not discussed or evaluated in the plenum. Our hypotheses were grounded in the assumption that increased knowledge would entail an increase in noticing: various authors assume that professional vision depends on the quantity and quality of the (domain-specific) knowledge (Seidel & Stürmer, 2014; Sherin & van Es, 2009; Sherin, 2007).

Obviously, even though knowledge, knowledge-based reasoning, and noticing were substantially interrelated, further subskills seem to play an important role in the ability of noticing, such as selective attention and situation awareness, which are further determined by contextual factors such as personal experiences or beliefs and thus lead to interindividual differences in information processing (Endsley, 1995; Sherin & van Es, 2009). A second explanation could refer to the video prompts used: They may have been too easy (in terms of test administration) to be able to assess changes or group differences (ceiling effects). Due to the scripted nature

of the video prompts, the classroom management issues were particularly salient. In a real setting, simultaneity and complexity of events are probably even greater. Moreover, the preservice teachers already had classroom management knowledge pre-intervention (about 50% correct answers in the MC-test-R, see Study 2). Looking at the results as a whole, recognizing relevant events may have been relatively easy for our participants, and the (differences in) knowledge acquisition may have been too small to substantially affect the ability to notice. These two possible explanations may provide interpretations of our results, particularly our finding that the training did not substantially affect noticing, in contrast to the results of other training studies (Santagata et al., 2007; Star & Strickland, 2008; van Es & Sherin, 2008).

8.1. Strengths and limitations

This study points to a number of both practical and theoretical issues that are important to teacher educators who are responsible for training preservice teachers. We implemented PBL in teacher education in classroom management and took a critical view of its potential to foster knowledge and professional vision. We therefore carried out two experimental studies that incorporated both quantitative and qualitative measures. In addition, we addressed Kirschner et al.'s (2006) critique that PBL should be adapted to the developmental needs of the learners, and is usually studied as a whole, instead of systematically varying single components of it (see Sweller et al., 2007).

Still, several shortcomings may limit the generalizability and scope of our findings. A first limitation is due to the research design applied here. In order to guarantee internal validity, we ran two consecutive studies in which we only varied a single aspect between the two groups: in Study 1, this was a single element of the learning environment, and in Study 2, it was the evaluation design to provide a valid assessment of change. However, this approach does not allow us to draw inferences about the (primarily) PBL nature of the treatment. We did not have a “real” control group such as a traditional PBL treatment in Study 1, or a further control group with a different treatment in Study 2. In addition, for curricular reasons, it was not possible to apply a pretest measure in Study 1 to determine the participants' preconditions. We therefore measured the effectiveness of the training only unspecifically. In future studies, to safely exclude exercise effects, a parallel test form should be created or a highly correlated measure of knowledge on classroom management should be applied. Although it seems valuable to further implement research designs like the one used here, we justified our decision with regard to internal validity. Still, the aforementioned flaws impede compatibility with previous research on PBL and the significance for the construction of suitable learning environments.

A second limitation should be noted with regard to the measurement occasions and the chosen instruments. First, research has demonstrated that participants in traditional PBL gain slightly less knowledge, but remember more of the acquired knowledge (Dochy et al., 2003). Due to accessibility (the preservice teachers finished their university exams just a few months later), we unfortunately had no opportunity to conduct a follow-up survey to assess the retention of the acquired knowledge and examine long-term effects on professional vision. Second, there were no established or available instruments available that could have been applied to measure our dependent variables accurately. Thus, the implemented measures were all newly developed. While correlations between the assessed constructs and external criteria were as expected, the reliability of the scores of the MC-tests were at the lower end and the assessment of professional vision was only qualitative in nature. Our attempt to quantify and condense the results of the video analyses may not have fully utilized the potential of the data and may also have limited the validity of the scores on professional vision. Especially with regard to knowledge-based reasoning the two theoretical facets, interpreting of observed incidents and developing appropriate strategies, were not measured and analyzed separately. In a future study, both facets should be gathered separately in order to obtain more information about students' competences. Moreover, the mere counting of mentioned incidents as an indicator for noticing probably did not cover this skill optimally, as it did not take into account selective attention and the relevance of incidents mentioned (Sherin, 2007). With regard to the latter, a criterion-oriented reference standard (e.g., an expert rating) would be desirable. It might also have been valuable to analyze the data with respect to content, e.g., whether participants in the CG overlooked or mentioned different events than the EG, or whether mentioned events became more relevant over time (see, e.g., Wolff et al., 2016). Although it goes beyond the scope of our study, it might have been useful to analyze the output from the process worksheets as well, particularly regarding the quantity and quality of generated alternative teacher strategies (see, e.g., Santagata & Angelici, 2010), which will be incorporated into our future research.

Finally, there may have been other reasons than the direct instruction in the EG that may have caused (no) group differences in Study 1. For example, CG students might have needed more time to become familiar with the literature and probably also may have wanted to have the opportunity to add literature they researched on their own. Although we tried to keep context features of the two groups as comparable as possible, which included keeping an even workload, we cannot rule out the possibility that other factors may have affected our dependent variables.

8.2. Conclusion and directions for future research

The current results provide support for the effectiveness of our training program, which was based on PBL and adapted to preservice teachers through elements of direct instruction. This is a first step towards systematic research on PBL in teacher education and instructional design. We demonstrated that it is worthwhile to explore the composition of learning environments in detail—in contrast to findings from previous research that focused on the comparison of entire instructional approaches, that is, PBL versus direct instruction curricula. It seems promising to keep employing experimental methodologies to gather empirical evidence and further elaborate on the impact of single elements of instructional settings (such as the amount of scaffolding, the composition of process worksheets, the problem design etc., see Kirschner et al., 2006) for different target groups and different target variables. An

innovative direction for improving the learning setting would be to further broaden sources of information for the problem-solving process, such as teacher or student interviews, which could add relevant information to appraise the classroom interactions and thus improve preservice teachers' emotional regulation, as well as their reflection and interpretation of upcoming classroom events.

To foster preservice teachers' professional vision of critical incidents in the classroom at the university level, more extensive training programs would also be viable. Curricula that combine the PBL cycle with a coherent set of opportunities to learn (e.g., in combination with microteaching or "video clubs" (Sherin & van Es, 2009; van Es, 2009), where students can start to explore the newly learned strategies in a shared environment under conditions of reduced complexity, Piwovar et al., 2013) could be studied in more detail. This should also include addressing a variety of other than cognitive outcomes, such as emotional regulation and performance in (simulated) teaching scenarios.

Declaration of interest

The authors declare that they have no conflict of interest in relation to this article.

Funding sources

This work was supported by the BMBF (Bundesministerium für Bildung und Forschung; Federal Ministry of Education and Research of Germany) [grant number 01JH1203].

Author note

Dr. Victoria Luise Barth, Department of Education and Psychology, Freie Universität Berlin, victoria.barth@fu-berlin.de. She is a research associate at the Department of Education and Psychology, Freie Universität Berlin (since 2012).

Dr. Valentina Piwovar, Department of Education and Psychology, Freie Universität Berlin, v.piwovar@fu-berlin.de. She was a research associate at the Department of Education and Psychology, Freie Universität Berlin (from 2009 to 2016).

Dr. Irina Rosa Kumschick, Department of Research and Development, Pädagogische Hochschule Luzern, irina-rosa.kumschick@phlu.ch. She is a research associate at the Department of Research and Development, Pädagogische Hochschule Luzern (since 2018).

Dr. Diemut Ophardt, School of Education, Technische Universität Berlin, ophardt@tu-berlin.de. She is the director of the School of Education (SETUB), Technische Universität Berlin (since 2015).

Prof. Dr. Felicitas Thiel, Department of Education and Psychology, Freie Universität Berlin, felicitas.thiel@fu-berlin.de. She is professor for School Pedagogy and School Improvement Research at Freie Universität Berlin (since 2006).

Acknowledgements

We would like to thank the preservice teachers, the university teachers, and the administrative staff involved for this research. We would also like to thank Dr. Franziska Pfitzner-Eden for her support in editing.

References

- Albanese, M. A., & Mitchell, S. (1993). Problem-based learning: A review of literature on its outcomes and implementation issues. *Academic Medicine*, 68, 52–81.
- Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*. <https://doi.org/10.1037/a0021017>.
- Aloe, A. M., Amo, L. C., & Shanahan, M. E. (2014). Classroom management self-efficacy and burnout: A multivariate meta-analysis. *Educational Psychology Review*. <https://doi.org/10.1007/s10648-013-9244-0>.
- Arthur, W., & Day, D. V. (1994). Development of a short form for the raven advanced progressive matrices test. *Educational and Psychological Measurement*, 54(2), 394–403.
- Barth, V. L. (2017). *Professionelle Wahrnehmung von Störungen im Unterricht [Professional vision of classroom disruptions]*. Wiesbaden: Springer VS.
- Barrows, H. S. (1996). Problem-based learning in medicine and beyond: A brief overview. In L. Wilkerson, & W. Gijsselaers (Eds.). *Bringing problem-based learning to higher education* (pp. 3–11). San Francisco: Jossey-Bass.
- Berliner, D. C. (2001). Learning about and learning from expert teachers. *International Journal of Educational Research*, 35(5), 463–482.
- Böttcher-Oschmann, F., Groß-Ophoff, J., & Thiel, F. (2018). *Validierung eines Fragebogens zur Erfassung studentischer Forschungskompetenzen über Selbsteinschätzungen. Ein Instrument zur Evaluation forschungsorientierter Lehr-Lernarrangements. [Validation of a questionnaire to assess university students' research competences – An instrument for evaluating research oriented teaching-learning arrangements.]* submitted.
- Brophy, J. E. (1988). Educating teachers about managing classrooms and students. *Learning and Teacher Education*. [https://doi.org/10.1016/0742-051X\(88\)90020-0](https://doi.org/10.1016/0742-051X(88)90020-0).
- Brophy, J. E. (2004). *Using video in teacher education. Advances in research on teaching, Vol. 10*. Amsterdam, London: JAI.
- Carter, K., Cushing, K., Sabers, D., Stein, P., & Berliner, D. C. (1988). Expert-novice differences in perceiving and processing visual classroom information. *Journal of Teacher Education*, 39(3), 25–31.
- Choi, H. J., & Yang, M. (2011). The effect of problem-based video instruction on student satisfaction, empathy, and learning achievement in the Korean teacher education context. *Higher Education*. <https://doi.org/10.1007/s10734-010-9403-x>.
- Christofferson, M., & Sullivan, A. L. (2015). Preservice teachers' classroom management training: A survey of self-reported training experiences, content coverage, and preparedness. *Psychology in the Schools*. <https://doi.org/10.1002/pits.21819>.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition). Hillsdale, NJ: Erlbaum.
- Dochy, F., Segers, M., van den Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: A meta-analysis. *Learning and Instruction*. [https://doi.org/10.1016/S0959-4752\(02\)00025-7](https://doi.org/10.1016/S0959-4752(02)00025-7).
- Doyle, W. (1984). How order is achieved in classrooms. An interim report. *Journal of Curriculum Studies*, 16(3), 259–277.
- Doyle, W. (1986). Classroom organization and management. In M. Wittrock (Ed.). *Handbook of research on teaching* (pp. 392–431). (3rd ed.). New York: Macmillan.
- Doyle, W. (2006). Ecological approaches to classroom management. In C. M. Evertson, & C. S. Weinstein (Eds.). *Handbook of classroom management* (pp. 97–125).

- Mahwah: Lawrence Erlbaum Associates.
- Edwards, S., & Hammer, M. (2006). Laura's story: Using problem based learning in early childhood and primary teacher education. *Learning and Teacher Education*, 22, 465–477.
- Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors the Journal of the Human Factors and Ergonomics Society*, 37(1), 32–64.
- Flaiss, J. L., & Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement*, 33(3), 613–619.
- Freeman, J., Simonsen, B., Briere, D. E., & MacSuga-Gage, A. S. (2014). Pre-Service teacher training in classroom management: A review of state accreditation policy and teacher preparation programs. *Teacher Education and Special Education*. <https://doi.org/10.1177/0888406413507002>.
- Gallagher, T. L. (2009). Teacher candidates' questions within the context of an educational psychology course. *Teaching Educational Psychology*, 5, 50–60.
- Gold, B., Förster, S., & Holodynski, M. (2013). Evaluation eines videobasierten Trainingsseminars zur Förderung der professionellen Wahrnehmung von Klassenführung im Grundschulunterricht [Evaluation of a video-based training to foster the professional vision of classroom management in elementary classrooms]. *Zeitschrift für Pädagogische Psychologie*, 27(3), 141–155.
- Golightly, A., & Raath, S. (2015). Problem-based learning to foster deep learning in preservice geography teacher education. *The Journal of Geography*. <https://doi.org/10.1080/00221341.2014.894110>.
- Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review*. <https://doi.org/10.1023/B:EDPR.0000034022.16470.f3>.
- Housner, L. D., & Griffey, D. C. (1985). Teacher cognition. Differences in planning and interactive decision making between experienced and inexperienced teachers. *Research Quarterly for Exercise and Sport*, 56(1), 45–53.
- Hushman, G., & Napper-Owen, G. (2011). Incorporating problem-based learning in physical education teacher education. *Journal of Physical Education Recreation & Dance*, 82, 17–23.
- Kirschner, P., Sweller, J., & Clark, R. (2006). Why minimal guidance during instruction does not work. *Educational Psychologist*. <https://doi.org/10.1207/s15326985ep4102.1>.
- Kounin, J. S. (1970). *Discipline and group management in classrooms*. New York, NY: Holt, Rinehart & Winston.
- Kulik, J. A., Kulik, C. C., & Bangert, R. L. (1984). Effects of practice on aptitude and achievement test scores. *American Educational Research Journal*, 21, 435–447.
- Kumschik, I. R., Piwowar, V., Ophardt, D., Barth, V. L., Krysmanski, K., & Thiel (2017). Optimierung einer videobasierten Lerngelegenheit im Problem Based Learning Format durch Cognitive Tools, Eine Interventionsstudie mit Lehramtsstudierenden [Optimizing a video-based learning environment in a problem based learning format using cognitive tools: An intervention study with preservice teachers.]. *Zeitschrift für Erziehungswissenschaft*, 20, 93–113. <https://doi.org/10.1007/s11618-017-0728-4>.
- Lachner, A., Jarodzka, H., & Nückles, M. (2016). What makes an expert teacher? Investigating teachers' professional vision and discourse abilities. *Instructional Science*. <https://doi.org/10.1007/s11251-016-9376-y> Advance online publication.
- LePage, P., Darling-Hammond, L., Akar, H., Gutierrez, C., Jenkins-Gunn, E., & Rosebrock, K. (2005). Classroom management. In L. Darling-Hammond, & J. Bransford (Eds.). *Preparing teachers for a changing world: What teachers should learn and be able to do* (pp. 327–357). San Francisco, CA: Jossey-Bass.
- Mayring, P. (2014). *Qualitative content analysis. Theoretical foundation, basic procedures and software solution*. Klagenfurt.
- Melnick, S., & Meister, D. (2008). A comparison of beginning and experienced teachers' concerns. *Educational Research Quarterly*, 31, 39–56.
- Moreno, R., & Valdez, A. (2007). Immediate and delayed effects of using a classroom case exemplar in teacher education: The role of presentation format. *Journal of Educational Psychology*. <https://doi.org/10.1037/0022-0663.99.1.194>.
- Nie, Y., & Lau, S. (2009). Complementary roles of care and behavioral control in classroom management: The self-determination theory perspective. *Contemporary Educational Psychology*. <https://doi.org/10.1016/j.cedpsych.2009.03.001>.
- Norman, G. R., & Schmidt, H. G. (1992). The psychological basis of problem-based learning: A review of the evidence. *Academic Medicine*, 67, 557–565.
- Ophardt, D., & Thiel, F. (2013). *Klassenmanagement, Ein Handbuch für Studium und Praxis [Classroom Management, A Handbook for Study and Practice]*. Stuttgart: Kohlhammer.
- Pfritzer-Eden, F., Thiel, F., & Horsley, J. (2014). An adapted measure of teacher self-efficacy for preservice teachers. *Zeitschrift für Pädagogische Psychologie*, 28, 1–10.
- Piwowar, V. (2013). *Konzeptualisierung, Erfassung und Entwicklung von Kompetenzen im Klassenmanagement. Dissertation zur Erlangung des akademischen Grades Doktor der Philosophie (Dr. phil.) [Conceptualisation, acquisition and development of competencies in classroom management. Dissertation for the academic degree of Doctor of Philosophy (Dr. phil.)]*. Berlin: Freie Universität Berlin.
- Piwowar, V., Thiel, F., & Ophardt, D. (2013). Training inservice teachers' competencies in classroom management. *Teaching and Teacher Education*, 30, 1–12. <https://doi.org/10.1016/j.tate.2012.09.007>.
- Piwowar, V., Barth, V. L., Ophardt, D., & Thiel, F. (2017). Evidence-based scripted videos on handling student misbehavior: The development and evaluation of video cases for teacher education. *Professional Development in Education*.
- Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, 6(2), 135–147.
- Santagata, R., & Angelici, G. (2010). Studying the impact of the Lesson Analysis Framework on preservice teachers' abilities to reflect on videos of classroom teaching. *Journal of Teacher Education*. <https://doi.org/10.1177/0022487110369555>.
- Santagata, R., Zannoni, C., & Stigler, J. W. (2007). The role of lesson analysis in pre-service teacher education: An empirical investigation of teacher learning from a virtual video-based field experience. *Journal of Mathematics Teacher Education*. <https://doi.org/10.1007/s10857-007-9029-9>.
- Schmidt, H. G. (1994). Problem-based learning: An introduction. *Instructional Science*, 22, 247–250.
- Seidel, T., & Stürmer, K. (2014). Modeling and measuring the structure of professional vision in preservice teachers. *American Educational Research Journal*. <https://doi.org/10.3102/0002831214531321>.
- Sherin, M. G. (2001). Developing a professional vision of classroom events. In T. L. Wood, B. S. Nelson, & J. Warfield (Eds.). *Beyond classical pedagogy. Teaching elementary school mathematics* (pp. S75–S93). Mahwah, N.J.: Erlbaum Associates.
- Sherin, M. G. (2007). The development of teachers' professional vision in video clubs. In R. Goldman, R. Pea, B. Barron, & S. Derry (Eds.). *Video research in the learning sciences* (pp. 383–395). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Sherin, M. G., & van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education*, 60(1), 20–37.
- Spiro, R. J., Feltoch, P. J., Jacobson, M. J., & Coulson, R. L. (1991). Cognitive flexibility, constructivism, and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains. *Educational Technology*, 31(5), 25–33.
- Star, J. R., & Strickland, S. K. (2008). Learning to observe. Using video to improve preservice mathematics teachers' ability to notice. *Journal of Mathematics Teacher Education*, 11, 107–125.
- Stough, L. M., & Montague, M. L. (2015). How teachers learn to be classroom managers. In E. Emmer (Ed.). *Handbook of classroom management: Research, practice, and contemporary issues* (pp. 446–458). Mahwah, NJ: Erlbaum.
- Stürmer, K., Seidel, T., & Holzberger, D. (2016). Intra-individual differences in developing professional vision. Preservice teachers' changes in the course of an innovative teacher education program. *Instructional Science*. <https://doi.org/10.1007/s11251-016-9373-1>.
- Sweller, J., Kirschner, P. A., & Clark, R. E. (2007). Why minimally guided teaching techniques do not work: A reply to commentaries. *Educational Psychologist*. <https://doi.org/10.1080/00461520701263426>.
- Thiel, F. (2016). *Interaktion im Unterricht. Ordnungsmechanismen und Störungsdynamiken [Interaction in the classroom. Order mechanisms and disturbance dynamics]*. Opladen & Tronoto: Verlag Barbara Budrich.
- Thiel, F., Richter, S. G., & Ophardt, D. (2012). Steuerung von Übergängen im Unterricht. Eine Experten-Novizen-Studie zum Klassenmanagement. [Management of transitions in the classroom – An expert-novice study about classroom management]. *Zeitschrift für Erziehungswissenschaft*, 4, 727–752.
- van den Bogert, N., van Bruggen, J., Kostons, D., & Jochems, W. (2014). First steps into understanding teachers' visual perception of classroom events. *Learning and Teacher Education*. <https://doi.org/10.1016/j.tate.2013.09.001>.

- van Es, E. A. (2009). Participants' roles in the context of a video club. *Journal of the Learning Sciences*, 18(1), 100–137.
- van Es, E. A., & Sherin, M. G. (2008). Mathematics teachers' "learning to notice" in the context of a video club. *Learning and Teacher Education*, 24, 244–276.
- van Tartwijk, J., & Hammerness, K. (2011). The neglected role of classroom management in teacher education. *Teaching Education*. <https://doi.org/10.1080/10476210.2011.567836>.
- Vernon, D. T., & Blake, R. L. (1993). Does problem-based learning work? A meta-analysis of evaluative research. *Academic Medicine*, 68, 550–563.
- Weinstein, C. S. (1998). "I want to be nice, but I have to be mean": Exploring prospective teachers' conceptions of caring and order. *Learning and Teacher Education*, 14, 153–163.
- Westerman, D. A. (1991). Expert and novice teacher decision making. *Journal of Teacher Education*. <https://doi.org/10.1177/002248719104200407>.
- Wettstein, A. (2013). Die Wahrnehmung sozialer Prozesse im Unterricht (The perception of social processes in classrooms). *Schweizerische Zeitschrift für Heilpädagogik*, 7/8, 5–13.
- Wijnia, L., Loyens, S. M. M., van Gog, T., Derous, E., & Schmidt, H. G. (2014). Is there a role for direct instruction in problem-based learning? Comparing student-constructed versus integrated model answers. *Learning and Instruction*. <https://doi.org/10.1016/j.learninstruc.2014.07.006>.
- Wolff, C. E., Jarodzka, H., van den Bogert, N., & Boshuizen, H. P. A. (2016). Teacher vision: Expert and novice teachers' perception of problematic classroom management scenes. *Instructional Science*. <https://doi.org/10.1007/s11251-016-9367-z>.
- Yoon, H., Woo, A., Treagust, D. F., & Chandrasegaran, A. (2014). The efficacy of problem-based learning in an analytical laboratory course for pre-service chemistry teachers. *International Journal of Science Education*, 36, 79–102.
- Zhang, M., Lundeberg, M. A., Koehler, M. J., & Eberhardt, J. (2011). Understanding affordances and challenges of three types of video for teacher professional development. *Learning and Teacher Education*. <https://doi.org/10.1016/j.tate.2010.09.015>.