



Metacognitive and multimedia support of experiments in inquiry learning for science teacher preparation

Till Bruckermann, Ellen Aschermann, André Bresges & Kirsten Schlüter

To cite this article: Till Bruckermann, Ellen Aschermann, André Bresges & Kirsten Schlüter (2017): Metacognitive and multimedia support of experiments in inquiry learning for science teacher preparation, International Journal of Science Education, DOI: [10.1080/09500693.2017.1301691](https://doi.org/10.1080/09500693.2017.1301691)

To link to this article: <http://dx.doi.org/10.1080/09500693.2017.1301691>



© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 19 Mar 2017.



Submit your article to this journal [↗](#)



Article views: 16



View related articles [↗](#)



View Crossmark data [↗](#)

Metacognitive and multimedia support of experiments in inquiry learning for science teacher preparation

Till Bruckermann^a, Ellen Aschermann^b, André Bresges^c and Kirsten Schlüter^a

^aInstitute for Biology Education, University of Cologne, Cologne, Germany; ^bEducational Psychology, University of Cologne, Cologne, Germany; ^cInstitute for Physics Education, University of Cologne, Cologne, Germany

ABSTRACT

Promoting preservice science teachers' experimentation competency is required to provide a basis for meaningful learning through experiments in schools. However, preservice teachers show difficulties when experimenting. Previous research revealed that cognitive scaffolding promotes experimentation competency by structuring the learning process, while metacognitive and multimedia support enhance reflection. However, these support measures have not yet been tested in combination. Therefore, we decided to use cognitive scaffolding to support students' experimental achievements and supplement it by metacognitive and multimedia scaffolds in the experimental groups. Our research question is to what extent supplementing cognitive support by metacognitive and multimedia scaffolding further promotes experimentation competency. The intervention has been applied in a two-factorial design to a two-month experimental course for 63 biology teacher students in their first bachelor year. Pre-post-test measured experimentation competency in a performance assessment. Preservice teachers worked in groups of four. Therefore, measurement took place at group level ($N=16$). Independent observers rated preservice teachers' group performance qualitatively on a theory-based system of categories. Afterwards, experimentation competency levels led to quantitative frequency analysis. The results reveal differing gains in experimentation competency but contrary to our hypotheses. Implications of combining scaffolding measures on promoting experimentation competency are discussed.

ARTICLE HISTORY

Received 3 April 2016

Accepted 27 February 2017

KEYWORDS

Self-regulation; video journal; experimentation competency; inquiry-based teaching; qualitative research; quasi-experimental research

Introduction

Inquiry skills do not just play a major role in students' education (e.g. National Science Teacher Association [NSTA], 2013; National Research Council [NRC], 2012) but in science teacher preparation, too. The NSTA (2012a) Preservice Science Standards emphasise that knowledge and practices of contemporary science are crucial for educating future science teachers. Both, knowing and practicing scientific methods are needed to develop knowledge through scientific inquiry for students in the classroom (NSTA, 2012b).

CONTACT Till Bruckermann  bruckermann@ipn.uni-kiel.de  Department of Biology Education, Leibniz Institute for Science and Mathematics Education, Kiel University, Kiel, Germany

© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Research has shown that preservice teachers' experiences in inquiry-based science teacher preparation facilitate the planning and teaching of inquiry-based lessons (Schwarz, 2009). Therefore, science teacher preparation should include inquiry-based university science courses (NSTA, 2003, p. 19). In the context of German science teacher preparation, the corresponding curriculum focuses on promoting knowledge and skills in hypotheses-led experimentation, as they are a part of subject-specific working and inquiry methods (Kultusministerkonferenz [KMK], 2010). Hypotheses-led experimentation constitutes one method of scientific inquiry and is characterised by testing a hypothesis on the causal relation of two variables while controlling for confounding variables (fair testing; Watson, Goldsworthy, & Wood-Robinson, 1999).

However, prospective science teachers showed deficiencies in scientific inquiry in general (Anderson, 2007) and when doing experiments in generating hypotheses and planning a referring experiment. Furthermore, they did not predict the outcome of their experiment and confounded variables (Hilfert-Rüppell et al., 2013). This gap between requirements on the one hand and preservice science teachers' deficiencies on the other has to be bridged in science teacher preparation by promoting experimentation competency. But experimentation is a cognitively complex problem-solving process (Klahr & Dunbar, 1988; Mayer, 2007), which is why scaffolding the learning process when doing experiments is necessary (Hmelo-Silver, Duncan, & Chinn, 2007). With this cognitive, metacognitive and multimedia measures of scaffolding can be distinguished: Multimedia tools seem promising to reduce cognitive load and facilitate reflection on the experimentation process (Bell, Urhahne, Schanze, & Ploetzner, 2010; Castek & Beach, 2013). Furthermore, applying methods of scientific inquiry requires planning, monitoring and self-reflecting learning strategies, in short metacognitive strategies. Therefore, encouraging prospective science teachers to use metacognitive control and regulation may supplement cognitive scaffolding in the experimentation process (Hilfert-Rüppell et al., 2013). This study aims at investigating the effects of metacognitive and multimedia scaffolding (while keeping cognitive support constant) on promoting experimentation competency of preservice teachers in a practical university course.

Theoretical framework

In this study, we tried to span the gap between the requirements for experimentation in science teacher preparation and the deficiencies in its promotion. We applied the theory of experimentation as a problem-solving process (Klahr & Dunbar, 1988; Mayer, 2007) to define the experimentation process, which is determined by problem-solving procedures in the context of knowledge. As competence features both knowledge and skills (Weinert, 2001), it is a central term in German standards (KMK, 2010). Therefore, preservice teachers' abilities were defined by experimentation competency. To promote experimentation competency, we considered cognitive load theory (CLT; Kirschner, Sweller, & Clark, 2006; Paas, Renkl, & Sweller, 2003) assuming that cognitive capacity is limited. This limitation might result in cognitive overload through inquiry learning, which is why we accounted for scaffolding in the experimentation process (Hmelo-Silver, 2006; Hmelo-Silver et al., 2007). Scaffolding focuses on cognitive (activity) and metacognitive (self-monitoring) strategies of learning (Davis & Linn, 2000). As experimentation competency comprises cognitive and metacognitive knowledge, we accounted for both types of

scaffolding. All measures we applied are scaffolds as it is not the question whether scaffolding is effective (Bannert, 2009) but in how far and for which outcomes different scaffolding environments are effective (Hmelo-Silver et al., 2007). Cognitive scaffolding of experiments in inquiry learning has proven to enhance students' experiments (Arnold, Kremer, & Mayer, 2014). Therefore, we provided all students with cognitive scaffolding (control) and varied the factor of supplemental metacognitive and multimedia scaffolding in a 2×2 factorial design to investigate whether further scaffolding adds further benefits.

For multimedia scaffolding, the cognitive theory of multimedia learning (CTML; Mayer, 2001) provided a framework of design principles to scaffold the experimentation process through the integration of multimedia tools (Castek & Beach, 2013). Furthermore, research on teacher students' experimentation competency reveals the necessity to include metacognitive guidance for planning, monitoring and self-reflecting on the experimentation process (Hilfert-Rüppell et al., 2013). Therefore, theory on metacognition in problem-solving procedures, which explains the necessity of regulating the learning process (Flavell, 1976; Zion, Michalsky, & Mevarech, 2012), framed approaches to support the experimentation process (Künsting, Kempf, & Wirth, 2013). The combination of both, metacognitive and multimedia scaffolding aimed for reciprocal effects, which means video production as well as self-reflection foster structured, thoughtful approaches when experimenting. Analysis of teacher students' experimentation competencies was based on performance assessment, which is a benchmark for measuring procedural knowledge on experimentation (Shavelson & Ruiz-Primo, 2005). It requires authentic tasks, observations on teacher students' performance and a scoring system. Therefore, we chose qualitative analysis (cf. Mayring, 2015) in order to structure the observed performance by categories and to measure it by a rating system we deduced from theory on experimentation as a problem-solving process (cf. Kremer, Specht, Urhahne, & Mayer, 2014; Mayer, 2007; Mayer, Grube, & Möller, 2008).

Experimentation competency

Competencies are dispositions that refer to a function, a context and a cognitive domain (Weinert, 2001). This means that experimentation competency is more specific than general problem-solving skills because it is applied to a context. Furthermore, experimentation competency is more than just isolated knowledge or procedural skills. Both, 'science practices' in the Next Generation Science Standards (NRC, 2012, p. 30) and competencies (KMK, 2010) emphasise the combined application of knowledge and skills.

In the sense of competencies in the German standards, domain-specific knowledge of and skills in hypotheses-led experimentation contribute to experimentation competency. As the German standards for biology teacher education require hypotheses-led experimentation (KMK, 2010), our research is focused on this method of scientific inquiry. Furthermore, Osborne, Collins, Ratcliffe, Millar, and Duschl (2003) emphasise the importance of experimentation in science. Hence, the present paper focuses on experimentation competency as a part of scientific competencies, admitting that other methods like observations and comparisons contribute equally to scientific progress (cf. NSTA, 2013).

Hypothesis-led experimentation constitutes one method of scientific inquiry and therefore, is linked to the competency of scientific reasoning (Mayer, 2007, p. 178). Hence, we apply the structural model of scientific reasoning by Mayer (2007) to

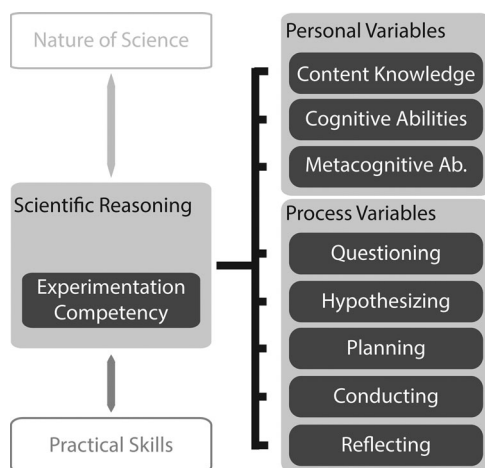


Figure 1. Model of scientific reasoning (adopted from Mayer, 2007).

define experimentation competency as depicted in Figure 1. It represents the two requirements, knowledge of (in Figure 1 referred to as personal variables) as well as skills (referred to as process variables) in hypotheses-led experimentation. Personal variables consider declarative content knowledge of concepts and procedural understanding contributing to experimentation competency. Furthermore, personal variables comprise cognitive as well as metacognitive abilities (Harms, 2007). The process variables include procedural knowledge of scientific inquiry like ‘formulating questions, generating hypotheses, designing and conducting investigations, and interpreting data’ (Arnold et al., 2014, p. 4; cf. Harwood, 2004; Klahr & Dunbar, 1988; Kremer et al., 2014; Meier & Mayer, 2012). Based on the process variables, Mayer et al. (2008; cf. Kremer et al., 2014) identified five inquiry levels, whose elaboration is represented by ascending roman numerals (I–V):

- (1) Investigation of a single factor: Experimental investigations comprise only one variable (Level I).
- (2) Investigation of a relationship: Two variables are causally linked and a conclusion can be drawn from observations (Level II).
- (3) Controlled investigation on the basis of conceptual knowledge: Considering control variables optimise the experimental design of two variables investigated (Level III).
- (4) Investigation of a generalised relationship: The hypotheses are generalisable and the experiment furthermore takes sample size, repetitive measurement and testing time into account. The discussion considers limitations of the findings (Level IV).
- (5) Investigative solution of a scientific problem: The research process is reflected regarding its accuracy and alternative conclusions are discussed (Level V).

Scaffolding of inquiry learning

There is discussion on the instructional quality of inquiry learning for promoting experimentation competency (Hmelo-Silver et al., 2007; Kirschner et al., 2006). Kirschner et al.

(2006) stated that inquiry learning is a ‘minimally guided approach’ (p. 99) and that this lack of guidance leads to cognitive overload. The authors base their argumentation on CLT (Paas et al., 2003), which assumes that space in working memory is limited. Working memory is overloaded by free exploration in inquiry tasks when learners do not possess adequate schemas to process the demands of the task (Kirschner et al., 2006). Hmelo-Silver et al. (2007) argue that inquiry learning is not minimally guided and ‘provide[s] extensive scaffolding and guidance to facilitate student learning’ (p. 1). Different measures of scaffolding can be distinguished. In this paper, we want to identify fruitful approaches to scaffold the experimentation process for promoting experimentation competency. Cognitive scaffolding structures the learning task and helps to accomplish the task by reducing cognitive load (Davis & Linn, 2000). Furthermore, metacognitive scaffolding enhances self-monitoring of the learning process on a more abstract level. Therefore, metacognitive support is a supplement to promote reflection on the experimentation process. Multimedia tools for making video clips scaffold the students by reducing cognitive load through documentation of observations, annotating them and enhancing metacognitive reflection (Quintana et al., 2004).

Cognitive support

Measures of cognitive scaffolding support students in applying cognitive learning strategies. Cognitive scaffolding is facilitated by displaying the structure, promoting self-explanations and giving expert guidance in the inquiry process (Hmelo-Silver, 2006). Structure, self-explanation and expert guidance are combined in incremental scaffolds, which could be used according to learners’ individual prerequisites (Arnold et al., 2014; Schmidt-Weigand, Hänze, & Wodzinski, 2009). Incremental scaffolds were administered in three steps:

- (1) ‘repeating definitions (e.g. what are dependent and independent variables)’ (Arnold et al., 2014, p. 22) gave a structure to the experimentation process by naming important steps such as determining variables,
- (2) ‘giving hints about the implementation’ (Arnold et al., 2014, p. 22) fostered students to explain to themselves how to implement these steps into their experiment and
- (3) ‘giving an exemplary answer’ (Arnold et al., 2014, p. 22) offered a possible solution, which could be used to check the accuracy of the self-explanation.

Arnold et al. (2014) provided help cards for prestructured experimental tasks, where upper-secondary students investigated everyday phenomena by formulating hypotheses, planning and conducting an experiment and making observations and interpreting their findings. The corresponding working material presented background information on the phenomenon. Scaffolding with help cards supported the students’ procedural knowledge with definitions, hints about implementation and exemplary answers for each step of the experiment. We used this cognitive approach for scaffolding experiments in the inquiry process of all control and experimental groups since in previous courses this measure has been proven to be an effective adaption from Arnold et al. (2014).

Metacognitive support

While cognitive support helps students to apply learning strategies, we intended metacognitive scaffolding to support students in planning, monitoring and self-reflecting the use of cognitive learning strategies, so that students have to interact and reflect on their experiments and overcome just technical aspects of experimentation (Gunstone & Champagne, 1990). This helps students to take control of their learning and becoming self-regulated through relying on their own planning, monitoring and evaluation of learning processes (Hofstein & Lunetta, 2004).

A model of self-regulated (Figure 2) learning mediates metacognitive learning strategies as it provides a meta-perspective on learning (Aschermann & Armbrüster, 2011). This model on self-regulation has proved to be efficient for scaffolding the use of metacognitive strategies because it reproduces the metacognitive activities of defining goals, planning, acting and evaluating the learning process and serves as a matrix for orientation on and internalisation of the learning process (Armbrüster, 2013). It corresponds to the inquiry process (generating questions and hypotheses, planning and conducting an experiment, interpreting data; Arnold et al., 2014). Different studies have shown that metacognitive scaffolding supports strategic learning processes (e.g. Bannert, 2003) and use of control-of-variables strategies in simulation-based discovery learning environments (Künsting et al., 2013).

In their study, Künsting et al. (2013) focused on metacognitive support of scientific discovery learning. They argued in favour of metacognitive support in the experimentation process because students have to regulate their use of general and domain-specific cognitive strategies. Their treatment utilised (1) an *introductory modelling* of metacognitive strategy use and furthermore (2) *prompting* verbally three times. The two measures can be distinguished as being (1) direct and (2) indirect: Direct measures should explicitly promote learning, whereas indirect measures scaffold learning in a more subtle way (Friedrich & Mandl, 1992). This results in direct measures being more effective but not that economic. Künsting et al. (2013) indicated a fostered



Figure 2. Cologne action-cycle model (Aschermann & Armbrüster, 2011; Bruckermann, Aschermann, Bresges, & Schlüter, 2015).

knowledge gain and more strategic experimenting when applying their treatment in a two-lesson computer-based learning environment. However, effectiveness of their economic metacognitive support has to be reassessed in real-life experimentation. Therefore, our study transferred their treatment of metacognitive support to experimentation in a university laboratory course.

Multimedia support

Computer tools are promising to support inquiry learning, when they are used as tools to implement scaffolding to learning (Bell et al., 2010; Quintana et al., 2004). The CTML (Mayer, 2001) provides a theoretical framework for learning with multimedia. It assumes that learners have to actively process information on a verbal and a visual channel, which are limited in capacity. Computer tools support the learning process by supporting routines in the experimental procedures (e.g. collecting data by photographs or measuring instruments). Thus, cognitive load in the experiment is reduced since space in the working memory is not occupied by procedures of documentation that could be automated. Furthermore, documentation in a movie provides multimodality (Castek & Beach, 2013). Multimodal representations help students to access information by reducing complexity (Bell et al., 2010). Apps for making movies of experiments (with a visual and sound track) facilitate multimodal transfer of observations and data and allow multiple representations. Multimodality addresses the possibility to represent and transfer data (e.g. observations) on several channels (Mayer, 2001). This transfer of data between verbal and visual channel helps to recognise the potentials of multimodal representations and urges the students from just interpreting data to convincing others by providing evidence (Castek & Beach, 2013).

Therefore, we intended mobile tablet computers to scaffold the experimentation process by producing video journals as an artefact of disciplinary strategies, automatising experiment documentation, using multimodal representations of observations and reflecting on the video journal (Quintana et al., 2004). Students use the integrated camera of tablet computers and an application for movie making (iMovie®) to document the experimentation process in so-called video journals. Journals require explaining how an experiment was planned and convincing the audience by presenting results. They support awareness of each step in the experimentation process and thus promote the understanding of principles of experimentation and scientific reasoning (Retzlaff-Fürst, 2013).

Research questions

Promotion of experimentation competency is an important goal in preservice science teacher preparation that can be attained by measures of (1) cognitive, (2) cognitive and metacognitive, (3) cognitive and multimedia or (4) cognitive, metacognitive and multimedia support (see Table 1). How these approaches work in the field of preservice teacher preparation has not been investigated yet, nor how they interact when applied in combination. Therefore, our research question is: Which of the four aforementioned approaches promotes students' learning of experimentation competency best?

Table 1. Comparison of three different approaches to promote experimentation competency (cf. Arnold et al., 2014; Künsting et al., 2013).

	Cognitive support (CG, EG1–3)	Metacognitive support (EG1 + 3)	Multimedia support (EG2 + 3)
Introduction to strategy use	<ul style="list-style-type: none"> Modelling of experimentation methodology: terms and concepts (e.g. hypotheses, variables, objectivity, reliability, validity) 	<ul style="list-style-type: none"> Modelling of metacognitive strategies: planning, monitoring and regulating learning strategies in the experimentation process 	<ul style="list-style-type: none"> Modelling of multimedia strategies for two applications: 1. iMovie[®] for using the camera and cutting movies, 2. Skitch[®] for annotating on pictures
Scaffolding the process	<ul style="list-style-type: none"> Help cards which repeat definitions, hint at the implementation and give exemplary answers to promote procedural knowledge of the experimentation process 	<ul style="list-style-type: none"> Prompting metacognitive strategy use by questions which provide consciousness, guidance and evaluation 	<ul style="list-style-type: none"> Using the camera for process documentation by taking pictures, annotating the process verbally (iMovie[®]) and visually (Skitch[®])
Reflection on the process	<ul style="list-style-type: none"> Based on a written lab journal 	<ul style="list-style-type: none"> Based on an action-cycle model 	<ul style="list-style-type: none"> Based on a video journal

Research methodology

The research question requires a 2×2 factorial quasi-experimental investigation of promoting experimentation competency. The two factors are metacognitive and multimedia support on inquiry learning. The control group (CG) got cognitive support with help cards. In experimental group 1 (EG 1), cognitive support with help cards was accompanied by an introductory modelling of metacognitive strategy use and further prompting (metacognitive support). In experimental group 2 (EG 2), cognitive support with help cards was accompanied by the production of video journals (multimedia support). In experimental group 3 (EG 3), cognitive support with help cards was accompanied by metacognitive and multimedia support as described above. We implemented the different treatments in a two-month practical course framed by pre- and post-performance assessment of the variables of experimentation competency. Performance assessments are the benchmark of measuring experimentation competency because they allow direct observations (Shavelson & Ruiz-Primo, 2005). Therefore, we decided to observe experimentation competency directly and analyse it qualitatively on the basis of a scoring system. Qualitative analysis of direct observations on experimentation competency comprised the description of levels and a frequency analysis of levels.

Study design

We chose a practical laboratory course with four parallel groups for implementation of the interventions during 10 weeks in the summer term of 2013 because it follows the learning goal of investigating phenomena by testing hypotheses on the causal relation of two variables while controlling for confounding variables. The knowledge base for investigating phenomena was prepared by a weekly 90-minute lecture on general biology by the course professor. In a weekly 90-minute practical course, student groups conducted experiments in the laboratory under the guidance of two graduate students teaching assistants

(instructors) for each course. The intervention spanned eight weeks of the practical course and two additional weeks for pre- and post-test. The students prepared each unit by reading the provided background information and subsequently testing their prior knowledge. Each unit started with a 5-minute introduction to the phenomenon and the research question to be investigated. We provided this research question to assure a comparable starting point. The involved researchers, who are from the fields of science education as well as educational psychology, designed the intervention and provided the instructors with all necessary materials. Regular meetings of the researchers and instructors guaranteed a comparable implementation of the intervention.

Cognitive inquiry support

For cognitive inquiry support, students (CG, EG 1–3) were provided with working materials which comprised a text with background information on the phenomenon and a derived research question. Affiliated tasks guided the students through the inquiry process. When a task overburdened students, they were able to access help cards. Help cards scaffolded the learning process by providing repetition of definitions, hints and exemplary solutions (cf. Arnold et al., 2014). The inquiry process had to be documented in a lab journal in order to allow reflection on it. At the end of each unit the lab journal was presented and reflected with all peers in the course.

Metacognitive inquiry support

For metacognitive support, we used an introductory modelling of 20 min in the first unit, which introduced metacognitive strategies by using the Cologne action-cycle model. Preservice teachers had to create a poster on which they emphasised the implications of the action-cycle for the experimentation process. Furthermore, the use of metacognitive strategies was prompted three times in each unit. The chosen prompts were unspecific and therefore provided flexibility on the situations they were used in (Bannert, 2009; Davis & Linn, 2000). Prompting was scheduled at the beginning, at the half and close at the end of a unit (Bannert, 2003). Unspecific but timed prompts should facilitate the recall of learned strategies of metacognitive support. Like Zion et al. (2012), we provided metacognitive guidance by three types of prompts. Metacognitive consciousness questions were provided before learning (e.g. Think of the action-cycle model.). Meta-guidance executive questions were provided during learning (e.g. Think of the phase of the action-cycle you are in and what is important about this phase.). At the end, questions prompted the evaluation of strategy use (e.g. Did you take enough time for each phase of the action-cycle model?).

Multimedia inquiry support

Video journaling considers the affordances of multimodality in two selected apps (movie making and picture annotating) for the experimentation processes and mediates scientific literacy (Castek & Beach, 2013). Documenting the inquiry process in a lab journal is an important tool for reflecting on it. It promotes understanding of the principles of experimentation and scientific reasoning (Retzlaff-Fürst, 2013). Therefore, students had to document their experiment in a video journal using apps on an iPad[®]. They were given an introductory instruction on the use of iMovie[®] (e.g. taking video clips, cutting them, adding soundtracks, blending in pictures) and Skitch[®] (e.g. taking pictures and annotating them). We provided the students with these two apps which mediated the production of video journals and therefore

multimodality. The immediate availability of a video journal on the tablet computer afforded the communication with peers by streaming it to a data projector. Thus, multimodality affordances in video journals promoted reflection on the experimentation process.

Sample

All preservice biology teachers, being qualified for schools of level 2 of the International Standard Classification of Education (ISCED) at the researchers' university in summer term 2013 participated in the investigation (representing a census of the bachelor programme, $N = 63$). Forty-six participants were female. Forty-four participants took part in an intensive biology course during school, 19 in a basic one. Only 14 participants had chosen biology in their final exam for high school graduation. Fifty-five of the participants were in their first year of a bachelor programme (eight were in a higher semester). Fifty-eight strived for the bachelor degree in biology education for lower secondary schools (ISCED level 2). $N = 63$ preservice teachers were equally assigned to the experimental groups considering only time preferences. All preservice teachers participated in the data collection. Each case for analysis considered a group of preservice teachers, which were compiled theory-based (Kelle & Kluge, 2010; Meier & Mayer, 2012) accounting for demographic factors (sex, age, semester) and previous knowledge (chemistry and biology) to achieve cooperative learning groups. Table 2 provides descriptive data on the control and experimental groups.

Ethics

All participants were informed about their participation in this research and all agreed on a written consent form. All data was anonymously processed.

Data collection

Performance assessment on experimentation competency

Different measurement procedures (e.g. paper-pencil-tests or performance assessment) result in measuring different traits of experimentation competency (Roberts & Gott,

Table 2. Overview on descriptive data on the control and experimental groups (NP = number of preservice teachers; f = female; m = male).

Control group ($N = 15$) Age $\bar{x} = 24.7$ (SD = 5.98); Sex % = 66.7 (female); 33.3 (male)			Experimental group 1 ($N = 16$) Age $\bar{x} = 20.9$ (SD = 1.71); Sex % = 73.3 (female); 26.7 (male)		
Case	NP	Sex	Case	NP	Sex
1	4	3 (f) 1 (m)	1	4	3 (f) 1 (m)
2	4	3 (f) 1 (m)	2	4	3 (f) 1 (m)
3	4	2 (f) 2 (m)	3	4	2 (f) 1 (m)
4	3	2 (f) 1 (m)	4	4	1 (missing) 3 (f) 1 (m)
Experimental Group 2 ($N = 16$) Age $\bar{x} = 21.3$ (SD = 2.85); Sex % = 81.3 (female); 18.8 (male)			Experimental Group 3 ($N = 16$) Age $\bar{x} = 21.6$ (SD = 3.79); Sex % = 75 (female); 25 (male)		
Case	NP	Sex	Case	NP	Sex
1	4	4 (f)	1	4	3 (f) 1 (m)
2	4	3 (f) 1 (m)	2	4	3 (f) 1 (m)
3	4	3 (f) 1 (m)	3	4	3 (f) 1 (m)
4	4	3 (f) 1 (m)	4	4	3 (f) 1 (m)

2003). This is due to the dependence of competencies on the context (Weinert, 2001). We decided to measure experimentation competency in a performance assessment as suggested by Shavelson and Ruiz-Primo (2005) to capture students' procedural knowledge as they conduct experiments. In order to measure the performance of preservice teachers in the inquiry process, three different components are needed for performance assessment: a task, a response format and a scoring system (Shavelson & Ruiz-Primo, 2005). As 'performance assessment scores are sensitive to the method used to assess performance' (Shavelson & Ruiz-Primo, 2005, p. 337), we chose a comparative investigation of plant cell reaction on different solutions of salt (sodium chloride) as the type of assessment. We used the same scenario for pre- and post-test because constructing equally demanding tasks for performance assessment is difficult due to task sampling variety. But since there was no ceiling effect, we conclude that using the same task did not affect the quality of measurement. The context of the task (osmosis) was addressed in a different experiment by the treatment. As competencies are contextual, the presentation of the phenomenon and the research question were embedded in a daily life situation (Why does salad get slack when the dressing is poured on it?). The time for solving the experimental task was 60 min. The preservice teachers were free to choose the procedures that were adequate to solve the task. The response format comprised 60 min videography of preservice teachers' performance on the experimental task, which was analysed through structuring and rating it by a system of categories on experimentation competency (scoring system). The data collection in the pre- and post-test framed the treatment and was not a part of the intervention. Groups' performance was videotaped by using one camera with additional microphone for each group in a non-participating observation. Based on the above-mentioned sample, we collected 16 videos of performance assessment each in the pre- and post-test. Unfortunately, one video of the post-test could not be analysed due to a failure of data processing and therefore was excluded from further analysis. Preservice teachers were informed that the assessment was not relevant for their marks on the practical course.

Data analysis

Performance assessment of experimentation competency requires direct observations which were videotaped in our study. Analysis of the video data was realised by a theory-based scoring system with categories and rating scales on preservice teachers' experimentation competency. Videos were analysed first to identify variables of the experimentation process by a deductive system of categories and second, to rate the elaboration of each category. Indicators of the experimentation competency levels were extracted from the video data for exemplification. Further frequency analysis of the ratings indicates changes from pre- to post-test.

Qualitative analysis

As suggested by Mayring (2000, 2015), we first deduced categories for analysis from the dimensions in theory (Arnold et al., 2014), coded and re-coded the material in iterative steps of revising the coding instructions until we finally coded and interpreted the results in a frequency analysis. The procedure is called scaled structuring (Mayring, 2015). The process variables of 'formulating questions, generating hypotheses, designing

and conducting investigations, and interpreting data' (Arnold et al., 2014, p. 4) were the dimensions of analysis (Mayring, 2015). Categories were derived from the theory-based dimensions by forming sub-competencies of experimentation competency (Kremer et al., 2014), for example, identifying variables (ref. to the dimension of designing investigation; see Table 3, second row), making observations and gathering data (ref. to the dimension of conducting investigations; see Table 3, third row and fourth row). Five different levels of competency were described for each category in ascending orders from level I to level V (cf. Kremer et al., 2014; Mayer et al., 2008; see Table 3). The empirically distinguishable competency levels were represented on a rating scale for each of the 10 categories. A manual on the system of categories guided the raters through the coding and rating process as suggested by Mayring (2015).

Analysis of preservice teachers' performance on experimentation competency considered 1740 min video data (900 min pre- and 840 min post-test video data), which was analysed by three independent observers (from the field of biology education) using MAXQDA 11 (VERBI, 2011). The observers were provided with a rating manual, which was improved at the beginning of the rating procedure by regularly communicating on the rating process following the guidelines of qualitative content analysis (Mayring, 2000, 2015). Two observers analysed 240 min of video data (approx. 14%) in duplication for calculating measurements on inter-observer reliability. Calculation of a reliability measure should be subsidiary to improving this measure for the rating procedure (cf.

Table 3. Levels of competency in scientific inquiry with reference to experiments (each category comprises the previous one; excerpt from manual by the authors; cf. Kremer et al., 2014; Mayer et al., 2008; Peeters, 2012).

Process variables (dimensions)	Sub-competencies (categories)	Inquiry skills (Levels)
'Generating hypotheses'	CC-1 <i>Generating hypotheses</i> [According to Mayer et al., 2008]	(5) Considering alternative hypotheses. (4) Generating testable and generalised hypotheses. (3) Generating testable hypotheses and explaining them on the basis of connectional knowledge. (2) Generating testable hypotheses and explaining them with analogies from everyday life. (1) Generating testable hypotheses. (0) Does not apply.
'Designing investigations'	CC-3 <i>Identifying variables</i> [According to Mayer et al., 2008]	(5) Considering accuracy of research methods. (4) Considering sample size, repetitive measurement, testing time. (3) Relating variables to each other and considering control variables. (2) Relating independent and dependent variables to each other. (1) Identifying a variable in an experimental setting. (0) Does not apply.
'Conducting investigations'	CC-8 <i>Making observations and gathering data</i> [According to Peeters, 2012]	(5) Observations and documentations are complete. (4) Observations are completely, but documentations scarcely made. (3) Observations and documentations are made, but scarcely. (2) Observations are scarcely made. (1) Observations are merely copied from other investigations. (0) Does not apply.
'Interpreting data'	CC-10 <i>Interpreting data</i> [According to Mayer et al., 2008; Peeters, 2012]	(5) Considering alternative conclusions and discussing limitations. (4) Reflecting conclusions from observation/data, concerning limitations and certainty aspects. (3) Relating variables to each other and considering control variables (2) Drawing conclusions from observations/data. (1) Reporting observations/data. (0) Does not apply.

Mayring, 2000; VERBI, 2014). Communication on the rating process served as an observer training for recalibration (Shavelson & Ruiz-Primo, 2005).

Frequency analysis

Frequency analysis is based on the ratings from qualitative analysis of the videos. Therefore, we accumulated the ratings on each level over all categories and plotted them in four diagrams for each group and test (pre- and post-test in Figure 3). Higher roman numerals represent higher competency levels. In order to describe changes in the competency levels, we checked whether statistically significant differences in learning gains occurred.

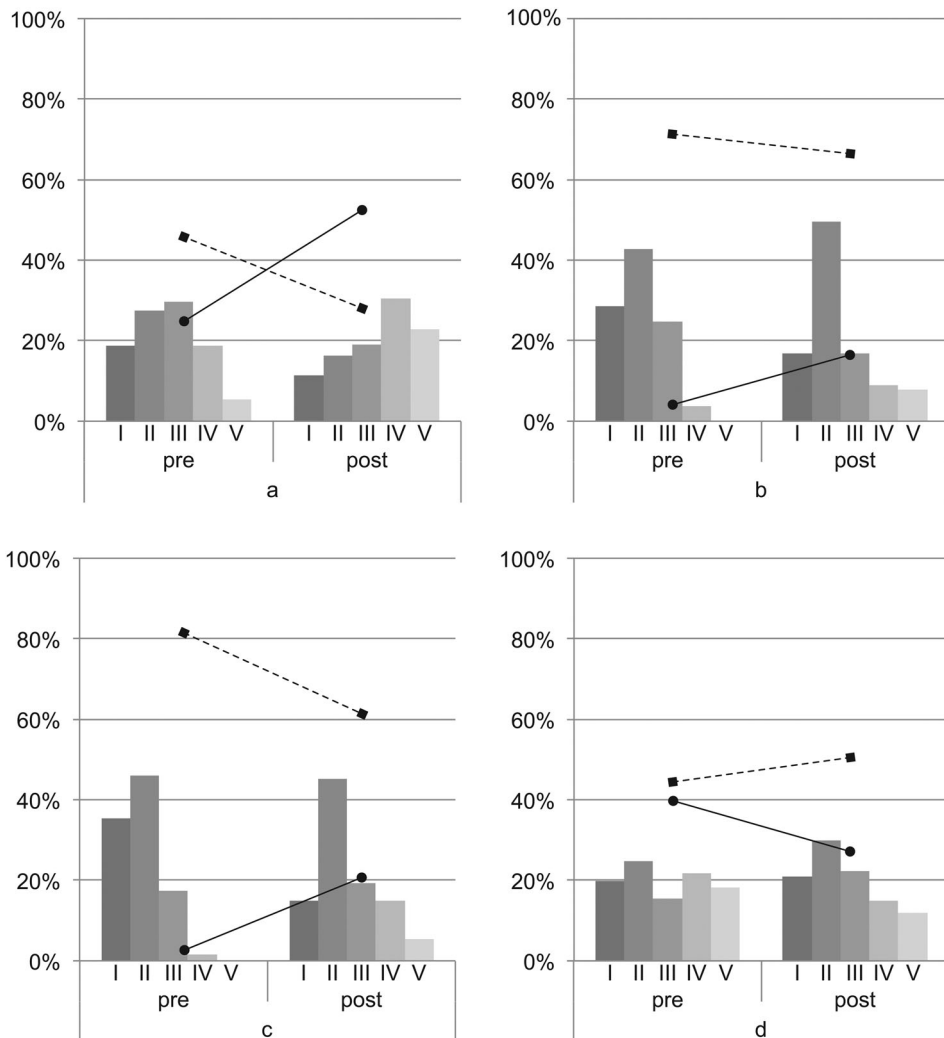


Figure 3. Frequency distribution of experimentation competency levels (higher roman numerals represent higher levels) for the (a) control group with cognitive support, (b) experimental group 1 with additional metacognitive support, (c) experimental group 2 with additional multimedia support and (d) experimental group 3 with additional metacognitive and multimedia support (accumulated ratings for Level IV/V: continuous line; Level I/II: broken line).

Learning gains were calculated by subtracting pre-test ratings from post-test ratings on each level to account for differences at the starting point. Afterwards, we compared the expected and the observed distributions of changes in the rating frequencies by a chi-square test to test whether the changes are meaningful. We would expect the same gains for all groups if there was no effect of the treatments. Effect sizes are expressed by ω^2 and evaluated according to Cohen (1988). As the chi-square test does not indicate the direction of the deviating distribution, the percentage frequencies are further inspected descriptively and used for discussion of effects. Tendencies on the changes in the rating frequency between pre- and post-test are indicated by broken lines for the accumulation of Levels I and II and continuous lines for Levels IV and V in Figure 3. While we expect the broken lines to decline from the pre- to the post-test, we expect the continuous lines to rise. For the different treatments, the lines should vary in their slopes. The more effective the treatment is the steeper the slope should be.

Results and discussion

Description of the sample and reliability

All categories and ratings, which were deduced from theory, are found throughout the video data. The following sections exemplarily show indicators of each experimentation competency level for the category *generating hypotheses* from the video data. The highest level (V) for *generating hypotheses* is indicated when preservice teachers are considering alternative hypotheses like they did in case 3 of the EG 3 in the post-test:

S1: The higher salinity, the higher water loss, the more lightweight the potatoes are. Rational: The high salinity imposes a concentration gradient from the potato towards the water. Therefore, water is detracted from the potato.

S2: The alternative hypothesis should be the other way round. (Post: EG3, Case 3, 00:18:25–00:19:14)

Level IV was reached by case 4 of the EG 3 in the pre-test by generating a testable and generalised hypothesis, which does not only consider two interacting factors but the concept:

S: The cell has a defined concentration of different salts and we have one solution of pure water without salt and two solutions with different concentrations of Na and Cl. And they are not present as molecules but as sole ions. And they get into the cell or they are in such concentration that the water gets out. Nuts, ions can't get in. They are too big. (Pre: EG3, Case 4, 00:21:21–00:21:51)

Level III considered testable hypotheses, which were explained on the basis of connectional knowledge like in case 4 of EG 1 in the post-test.

S: The higher concentrated the solution of salt is, the more lightweight the potatoes get. (Post: EG1, Case 4, 00:09:12–00:09:27)

The lowest level (I) for generating hypotheses is indicated when preservice teachers generated a testable hypotheses like in case 4 of EG1 in the pre-test:

S: Our hypothesis is that more salt dries out the potatoes. (Pre: EG1, Case 4, 00:04:26–00:04:53)

The overall sample exhibited all levels of experimentation competency in pre- and post-test (pre: 572 ratings = 100%; post: 402 ratings = 100%). In the pre-test almost 60% of ratings were on Levels I and II indicating *investigations of only one single factor* (Level I) or a *relationship* (Level II). Level III, which means a *controlled investigation based on conceptual knowledge*, comprised 22% of ratings. Nevertheless, 13% of the ratings indicated an *investigation of a generalised problem* (Level IV), while 6% of the ratings were on the highest level of experimentation competency, which is *solving a scientific problem in an investigative manner* (Level V). In comparison to the pre-test, the number of ratings on Levels IV and V increased in the post-test, while the number of ratings on Levels III and I decreased. Ratings on Level II stagnated.

According to Mayring (2015, p. 125), qualitative content analysis considers further criteria for reliability (e.g. guidance by standards and documentation of the process) and validity besides coefficients (e.g. closeness to the subject, validation of interpretations and triangulation). We assured reliability of the analysis by setting up a theoretical framework (Klahr & Dunbar, 1988; Mayer, 2007) and following the standards according to Mayring (2000, 2015). A manual with the description of the process and definitions on the units of analysis, dimensions, categories and the rating scale documented the process and accompanied the analysis for reliable measurement. We accounted for the promotion of experimentation competency as required by preservice teacher education standards (KMK, 2010) and thus provided curricular validity. Although the analysis proceeded theory-based, application of categories might be subjective. Therefore, two independent observers, who were previously not involved in the research process, validated the inter-subjectivity of interpretations. We calculated observer agreement for two independent observers on the categories and ratings (on approx. 14% of the material), which turned out to be moderate ($\kappa = .58$ Cohen's kappa; Landis & Koch, 1977) but acceptable (cf. Greve & Wentura, 1997) for each pre- and post-test. Regular meetings and communication improved the manual on the system of categories and served as a recalibration of the observers during the analysis.

Comparisons on the levels of experimentation competency

The frequency distribution analysis shows increases and decreases of ratings on experimentation competency levels (ascending roman numerals) from pre- to post-test for CG (Figure 3(a)), EG 1 (Figure 3(b)), EG 2 (Figure 3(c)) and EG 3 (Figure 3(d)). On a statistical basis, we observe that learning gains differed between all groups, with $\chi^2(12) = 51.31, p < .05, \omega^2 = 0.17$ ('medium' effect size; Cohen, 1988). Post-hoc descriptive analysis of the frequencies reveals that only EG 3 with metacognitive and multimedia support showed decreases in the higher (IV and V) and increases in the lower levels (I–II), while the other three groups (CG, EG 1, EG 2) showed an increase in the higher (IV and V) and a decrease in the lower levels (I–II).

Changes in experimentation competency in CG (cognitive support)

When being compared to the other samples, the control group showed the greatest increases in competency levels (29% on Levels IV and V). Both upper levels (IV and V) amounted to more than half of all ratings (53%) on this group in the post-test. They gained 18% more ratings on Level V than in the pre-test, which is indicated by the solid line in Figure 3(a). This indicates a change towards the *investigative solution of*

scientific problems (Level V), whereas ratings on Level I decreased by 8% and on Level II by 11% (broken line in [Figure 3\(a\)](#)). Thus, the results show that cognitive support promoted experimentation competency. The basic cognitive support with help cards promoted reflection on the experimentation process and thus, seems to have scaffolded students' learning process: (a) structuring the experimentation process, (b) asking students for self-explanations about implementation and (c) giving expert guidance by an exemplary answer (cf. Arnold et al., 2014; Hmelo-Silver et al., 2007). Furthermore, journals of their investigation were a basis for a concluding discussion to promote the reflection on scientific reasoning. The findings by Arnold (2015) support this explanation where help cards reduced cognitive load during experimentation.

Changes in experimentation competency in EG 1 (cognitive and metacognitive support)

[Figure 3\(b\)](#) shows that the ratings on the upper experimentation competency levels (IV and V) increased for the group with metacognitive support (EG 1) by 13% from pre- to post-test while ratings on Levels I and III decreased. The increases were by 16% smaller than in the control group. Furthermore, ratings on Levels II increased, too, indicating gains on a lower level. Students with additional metacognitive support (EG 1) had no higher gains in experimentation competency than the control group, which indicates that additional metacognitive support did not outperform the cognitive support of the control group. In this experimental group, teaching was facilitated by additional metacognitive support to promote reflection on learning strategies in the experimentation process. Regarding the results, students might have lacked the abilities to employ metacognitive strategies as they were in their first year of bachelor. Thus, a conflict between the use of cognitive and metacognitive strategies (Clark, 1987) hindered learning in comparison to the control group. According to the results of Kempf and Künsting (2013), we recommend testing the effectiveness of metacognitive strategy use with more experienced learners in inquiry-based experimentation. However, we want to discuss in how far the cognitive scaffolds overlapped metacognitive scaffolding: cognitive scaffolding might have been metacognitive in its nature, too, as it promoted self-explanation and group discussion and therefore, enhanced explicit self-reflection in the experimentation process. Further metacognitive scaffolds were more general than cognitive scaffolds in relating to the learning process (e.g. Think of the phase of the action-cycle you are in and what is important about this phase.). Therefore, metacognitive scaffolds might act as a weaker support strategy.

Changes in experimentation competency in EG2 (cognitive and multimedia support)

The results on EG 2 with additional multimedia support are pointing towards no higher gains in experimentation competency compared to the control group with cognitive support. Although EG 2 gains 19% more ratings on Levels IV and V of experimentation competency as depicted in [Figure 3\(c\)](#), it has by 10% inferior rating gains in comparison to the control group (which had 29% additional rating gains). Smaller gains on experimentation competency due to multimedia support can be explained by CLT: Since the use of multimedia to document the experimentation process in video journals is new to the pre-service teachers, it imposes extraneous cognitive load on them. Therefore, multimedia support might have hindered further promotion of experimentation competency when being compared to the control group. Mayer and Moreno (2003) already showed

detrimental effects of multimedia learning when the processing of information overemphasises one channel (dual-channel assumption) and therefore results in cognitive overload. These overload scenarios can be applied to video journaling, for example, presenting and naming materials visually (which means showing a sequence of redundant images and text blocks simultaneously) causes cognitive overload as essential processing demands are accompanied by incidental processing demands due to confusing material (Mayer & Moreno, 2003, p. 46). Therefore, load-reducing methods may help preservice teachers to profit from video journaling because they provide guidance on how to produce video journals. However, this requires integrating the promotion of multimedia strategies in the practical course (Scheiter, Schubert, Gerjets, & Stalbovs, 2014), which might overload course content.

Changes in experimentation competency in EG3 (combining cognitive, metacognitive and multimedia support)

EG 3 (see Figure 3(d)) showed a quite different development of experimentation competency compared to the other groups. Although all other groups (CG, EG 1 and EG 2) are gaining higher competency levels, EG 3 loses higher levels (IV and V) and gains in the lower levels (I–III). Combining metacognitive and multimedia support did not outperform basic cognitive support. It seems to be even worse: Cognitive overload caused by the video journals was complemented by the detrimental effect of two competing strategies, cognitive support and metacognitive support.

Limitations

However, we have to admit the following limitations of this study: First, the conclusions are drawn on the basis of a small sample ($N_{\text{participants}} = 63$; $N_{\text{groups}} = 16$), which still represents all enrolled preservice teachers for ISCED level 2 in their first bachelor year at the reviewed university. Furthermore, statistical analysis by chi-square test is limited to a general comparison of all groups. Post-hoc descriptive analysis does not tell which single learning gains were significant. The second limitation of this study concerns the emphasis of summative evaluation. Further research should additionally take the process of learning into account and analyse the video journals compared to written lab journals created by students during the treatment. Especially comparative investigations of the lab journals can reveal further learning outcomes such as improvements in using technical language or making movies. Third, implementing the treatment into a regular university course during a time span of two months required accounting for confounding variables. To account for a quasi-experimental design, participants were assigned just regarding their time preferences. To assure a comparable implementation of the intervention, we provided the different instructors with detailed manuals for applying the treatment. Although we accounted for different confounding variables in a quasi-experimental design, we investigated the treatment under conditions that are typical for university courses and therefore increased ecological validity (Brewer, 2000).

Conclusions and implications

The focus of this study was to investigate a measure to bridge the gap of requirements (NRC, 2012; KMK, 2010) and the deficiencies of teacher students in scientific inquiry

(Anderson, 2007; Hilfert-Rüppell et al., 2013). We do this at the example of a practical course with inquiry-based experiments to promote experimentation competency of pre-service teacher students in their first year of a bachelor programme because inquiry-based teacher education is a prerequisite for inquiry in schools (Capps & Crawford, 2013). As inquiry-based learning can be facilitated by different measures of scaffolding, we compared sole cognitive with additional metacognitive and multimedia scaffolds to identify their influence on experimentation competency as learning outcome. Although the sample was small and requires replication of findings at other universities, it gave insights on different opportunities for promoting experimentation competency at this exemplary university. Implementing the treatment in the running teaching programme provided high ecological validity, while the quasi-experimental design accounted for confounding variables. Our findings align with previous findings (Arnold et al., 2014; Schmidt-Weigand et al., 2009) that cognitive support with help cards provides scaffolding to the experimentation process and promotes experimentation competency by promoting self-explanations and giving expert examples as reference. Furthermore, we have learned that students can lack the abilities to profit from metacognitive support resulting in a conflict of cognitive and metacognitive strategy use. However, cognitive scaffolds that require self-explanation enhance explicit reflection on the experimentation process. Similar effects have been observed for multimedia support: Even though multimedia tools provide scaffolding by explicating disciplinary reasoning strategies (Hmelo-Silver et al., 2007; Quintana et al., 2004), multimedia tools can induce extraneous cognitive load, when their application requires cognitive capacity that is additional to the task (Mayer & Moreno, 2003). Hence preservice teachers suffered from cognitive overload and gained not as much in experimentation competency as preservice teachers with sole cognitive support did. It is important to recognise that learning with multimedia has to account for cognitive load and depends on mastering multimedia strategies. Strategy training should be integrated in regular instruction as it is a prerequisite for learning from multimedia (Scheiter et al., 2014). This influences future learning goals, which cannot only be limited to experimentation competency but have to include multimedia strategies when working with video journals. Thus, combining different measures of scaffolding does not necessarily enhance learning outcomes. We conclude, that metacognitive scaffolding does not outperform cognitive scaffolding of inquiry learning when the cognitive condition promotes explicit reflection on the experimentation process (Duschl & Grandy, 2013). Incremental scaffolds providing structure, self-explanations and expert guidance are an effective measure to promote experimentation competency by including explicit reflection.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

- Anderson, R. D. (2007). Inquiry as an organizing theme for science education. In S. K. Abell, & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 807–830). Mahwah, NJ: Erlbaum.
- Armbrüster, C. (2013). *Lernprozesse vielfältig gestalten: Entwicklung, Durchführung und Evaluation eines Modells zur Förderung des selbstregulierten Lernens* [Creating divers learning processes:

- development, implementation and evaluation of a model on promotion of self-regulated learning]. Hamburg: Verlag Dr. Kovač.
- Arnold, J. (2015). *Die Wirksamkeit von Lernunterstützungen beim Forschenden Lernen. Eine Interventionsstudie zur Förderung des Wissenschaftlichen Denkens in der gymnasialen Oberstufe* [Efficacy of learning support on inquiry learning. Intervention study on promotion of scientific reasoning in secondary schools]. Berlin: Logos.
- Arnold, J. C., Kremer, K., & Mayer, J. (2014). Understanding students' experiments – what kind of support do they need in inquiry tasks? *International Journal of Science Education*, 36(16), 2719–2749. doi:10.1080/09500693.2014.930209
- Aschermann, A., & Armbrüster, C. (2011). *Get involved – Persönliche Kompetenzen erkennen und fördern* [Identifying and promoting personal competencies]. Abschlussbericht. Köln: Universität zu Köln.
- Bannert, M. (2003). Effekte metakognitiver Lernhilfen auf den Wissenserwerb in vernetzten Lernumgebungen [Effects of Metacognitive Help on Knowledge Acquisition in Web-based Learning Environments]. *Zeitschrift für Pädagogische Psychologie*, 17(1), 13–25.
- Bannert, M. (2009). Promoting self-regulated learning through prompts. *Zeitschrift für Pädagogische Psychologie*, 23(2), 139–145.
- Bell, T., Urhahne, D., Schanze, S., & Ploetzner, R. (2010). Collaborative inquiry learning: Models, tools, and challenges. *International Journal of Science Education*, 32(3), 349–377. doi:10.1080/09500690802582241
- Brewer, M. B. (2000). Research design and issues of validity. In H. T. Reis, & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 3–16). New York, NY: Cambridge University Press.
- Bruckermann, T., Aschermann, E., Bresges, A., & Schlüter, K. (2015). Promoting teacher trainees' inquiry skills by self-regulation and tablets. In K. Maaß, G. Törner, D. Wernisch, E. Schäfer, & K. Reitz-Koncebovski (Eds.), *Educating the educators: International approaches to scaling-up professional development in mathematics and science education* (pp. 135–146). Münster: WTM-Verlag.
- Capps, D. K., & Crawford, B. A. (2013). Inquiry-based professional development: What does it take to support teachers in learning about inquiry and nature of science? *International Journal of Science Education*, 35(12), 1947–1978.
- Castek, J., & Beach, R. (2013). Using apps to support disciplinary literacy and science learning. *Journal of Adolescent & Adult Literacy*, 56(17), 554–564.
- Clark, R. E. (1987, April). *When teaching kills learning: Types of mathemathantic effects*. Paper presented at the Annual Meeting of the American Educational Research Association. Washington, DC, USA.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.
- Davis, E. A., & Linn, M. C. (2000). Scaffolding students' knowledge integration: Prompts for reflection in KIE. *International Journal of Science Education*, 22(8), 819–837. doi:10.1080/095006900412293
- Duschl, R. A., & Grandy, R. (2013). Two views about explicitly teaching nature of science. *Science & Education*, 22(9), 2109–2139. doi:10.1007/s11191-012-9539-4
- Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *Perspectives on the development of memory and cognition* (pp. 231–235). Hillsdale, NJ: Lawrence Erlbaum.
- Friedrich, H. F., & Mandl, H. (1992). Lern- und Denkstrategien – ein Problemaufriß [Learning and thinking strategies – Outlining the problem]. In H. Mandl, & H. F. Friedrich (Eds.), *Lern- und Denkstrategien* (pp. 3–54). Göttingen: Hogrefe.
- Greve, W., & Wentura, D. (1997). *Wissenschaftliche Beobachtung: Eine Einführung* [Scientific observations: an introduction]. Weinheim: Beltz.
- Gunstone, R. F., & Champagne, A. B. (1990). Promoting conceptual change in the laboratory. In E. Hegarty-Hazel (Eds.), *The student laboratory and the science curriculum* (pp. 159–182). London: Routledge.

- Harms, U. (2007). Theoretische Ansätze zur Metakognition [Theoretical approaches to metacognition]. In D. Krüger, & H. Vogt (Eds.), *Theorien in der biologiedidaktischen Forschung* (pp. 177–186). Berlin: Springer.
- Harwood, W. S. (2004). A new model for inquiry – Is the scientific method dead? *Journal of College Science Teaching*, 33(7), 29–33.
- Hilfert-Rüppell, D., Looß, M., Klingenberg, K., Eghtessad, A., Höner, K., Müller, R., ... Pietzner, V. (2013). Scientific reasoning of prospective science teachers in designing a biological experiment. *Lehrerbildung auf dem Prüfstand*, 6(1), 135–177.
- Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, 42(2), 99–107.
- Hmelo-Silver, C. E. (2006). Design principles for scaffolding technology-based inquiry. In A. M. O'Donnell, C. E. Hmelo-Silver, & G. Erkens (Eds.), *Collaborative learning, reasoning, and technology* (pp. 147–170). Mahwah, NJ.
- Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education*, 88, 28–54. doi:10.1002/scs.10106
- Kelle, U., & Kluge, S. (2010). *Vom Einzelfall zum Typus: Fallvergleich und Fallkontrastierung in der qualitativen Sozialforschung* [From a single case to a type: Case comparison and contrasting]. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kempf, J., & Künsting, J. (2013). Wirksamkeit metakognitiver Lernhilfen bei entdeckendem Lernen durch Experimentieren in der Sekundarstufe I [Effectiveness of metacognitive Support on Discovery Learning through Experimentation in the Lower Secondary]. *Psychologie in Erziehung und Unterricht*, 60(4), 267–281.
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41(2), 75–86.
- Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science*, 12(1), 1–48.
- Kremer, K., Specht, C., Urhahne, D., & Mayer, J. (2014). The relationship in biology between the nature of science and scientific inquiry. *Journal of Biological Education*, 48(1), 1–8.
- Kultusministerkonferenz (KMK, Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany). (2010). *Laendergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und die Fachdidaktiken in der Lehrerbildung* [State common curricular requirements on science and education in teacher education]. Retrieved December 13, 2012, from http://www.akkreditierungsrat.de/fileadmin/Seiteninhalte/KMK/Vorgaben/KMK_Lehrerbildung_inhaltliche_Anforderungen_aktuell.pdf
- Künsting, J., Kempf, J., & Wirth, J. (2013). Enhancing scientific discovery learning through metacognitive support. *Contemporary Educational Psychology*, 38, 349–360.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- Mayer, R. E. (2001). *Multimedia learning*. New York, NY: Cambridge University Press.
- Mayer, J., Grube, C., & Möller, A. (2008). Kompetenzmodell naturwissenschaftlicher Erkenntnisgewinnung [Competency model of scientific inquiry]. In U. Harms & A. Sandmann (Eds.), *Lehr- und Lernforschung in der Biologiedidaktik* (Vol. 3, pp. 63–79). Innsbruck: Studienverlag.
- Mayer, J. (2007). Erkenntnisgewinnung als wissenschaftliches Problemlösen [Inquiry as scientific reasoning]. In D. Krüger, & H. Vogt (Eds.), *Theorien in der biologiedidaktischen Forschung* (pp. 177–186). Berlin: Springer.
- Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 43–52. doi:10.1207/S15326985EP3801_6
- Mayring, P. (2000). Qualitative content analysis [28 paragraphs]. *Forum Qualitative Sozialforschung*, 1(2), 1–10. Art. 20. Retrieved December 1, 2015, from <http://nbn-resolving.de/urn:nbn:de:0114-fqs0002204>

- Mayring, P. (2015). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* [Qualitative Content Analysis. Foundation and Techniques]. Weinheim: Beltz.
- Meier, M., & Mayer, J. (2012). Experimentierkompetenz praktisch erfassen – Entwicklung und Validierung eines anwendungsbezogenen Aufgabendesigns [Assessing experimentation competency practically – Development and validation of an applied task]. In U. Harms, & F. X. Bogner (Eds.), *Lehr- und Lernforschung in der Biologiedidaktik* (Vol. 5, pp. 81–98). Innsbruck: Studien Verlag.
- National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.
- National Science Teachers Association. (2003). *Standards for science Teacher preparation, Revised 2003*. Retrieved January 8, 2015, from <http://www.winthrop.edu/uploadedFiles/cas/biology/NSTASTandards2003.pdf>
- National Science Teachers Association. (2012a). *Preservice science standards*. Retrieved November 11, 2015, from <http://www.nsta.org/preservice/docs/2012NSTAPreserviceScienceStandards.pdf>
- National Science Teachers Association. (2012b). *Knowledge base supporting the preservice science standards*. Retrieved November 11, 2015, from <http://www.nsta.org/preservice/docs/KnowledgeBaseSupporting2012Standards.pdf>
- National Science Teachers Association. (2013). *Next generation science standards: For states, by states*. Retrieved February 28, 2016, from <http://www.nextgenscience.org/next-generation-science-standards>
- Osborne, J., Collins, S., Ratcliffe, M., Millar, R., & Duschl, R. (2003). What ‘ideas-about-science’ should be taught in school science? A Delphi study of the expert community. *Journal of Research in Science Teaching*, 40(7), 692–720.
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38(1), 1–4.
- Peeters, W. (2012, March). *Evaluating research competences: Methods and strategies*. Poster presented at training session “5: Deepening the understanding of inquiry in natural sciences” Paris, France.
- Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., ... Soloway E. (2004). A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences*, 13, 337–386.
- Retzlaff-Fürst, C. (2013). Protokollieren, Zeichnen und Mathematisieren [Journalizing, drafting and mathematizing]. In H. Gropengießer, U. Harms, & U. Kattmann (Eds.), *Fachdidaktik Biologie* (pp. 312–324). Hallbergmoos: Aulis Verlag.
- Roberts, R., & Gott, R. (2003). Assessment of biology investigations. *Journal of Biological Education*, 37(3), 114–121.
- Scheiter, K., Schubert, C., Gerjets, P., & Stalbovs, K. (2014). Does a strategy training foster students’ ability to learn from multimedia? *The Journal of Experimental Education*, 83(2), 266–289.
- Schmidt-Weigand, F., Hänze, M., & Wodzinski, R. (2009). Complex problem solving and worked examples: The role of prompting strategic behavior and fading-in solution steps. *Zeitschrift für pädagogische Psychologie*, 23(2), 129–138.
- Schwarz, C. (2009). Developing preservice elementary teachers’ knowledge and practices through modeling-centered scientific inquiry. *Science Education*, 93, 720–744.
- Shavelson, R. I., & Ruiz-Primo, M. (2005). On the psychometrics of assessing science understanding. In J. J. Mintzes, J. H. Wandersee, & J. D. Novak (Eds.), *Assessing science understanding. Educational psychology* (pp. 303–341). Burlington: Academic Press.
- VERBI Software. Consult. Sozialforschung GmbH. (2011). Software for qualitative data analysis Version 11.
- VERBI Software. Consult. Sozialforschung GmbH. (2014). *Manual on MAXQDA*. Retrieved January 29, 2016, from <http://www.maxqda.de/support/hilfe/maxqda-12/#!Documents/willkommen.htm>
- Watson, R., Goldsworthy, A., & Wood-Robinson, V. (1999). What is not fair with investigations? *School Science Review*, 80(292), 101–106.

- Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen, & L. H. Salganik (Eds.), *Defining and selecting key competencies* (pp. 45–65). Goettingen: Hogrefe.
- Zion, M., Michalsky, M., & Mevarech, Z. R. (2012). The effects of metacognitive instruction embedded within an asynchronous learning network on scientific inquiry skills. *International Journal of Science Education*, 27(8), 957–983. doi:10.1080/0950069050006862