TYPE Original Research
PUBLISHED 18 March 2024
DOI 10.3389/feduc.2024.1293516

OPEN ACCESS

EDITED BY
Niwat Srisawasdi,
Khon Kaen University, Thailand

REVIEWED BY
Stephanie M. Gardner,
Purdue University, United States
Katrin Sommer,
Ruhr University Bochum, Germany

*CORRESPONDENCE
Miriam Lechner
✉ miriam.degner@tum.de

RECEIVED 13 September 2023
ACCEPTED 29 February 2024
PUBLISHED 18 March 2024

CITATION
Lechner M, Moser S, Pander J, Geist J and Lewalter D (2024) Learning scientific observation with worked examples in a digital learning environment.
Front. Educ. 9:1293516.
doi: 10.3389/feduc.2024.1293516

Learning scientific observation with worked examples in a digital learning environment

Miriam Lechner[1]*, Stephanie Moser[1], Joachim Pander[2], Juergen Geist[2] and Doris Lewalter[1]

[1]Department Educational Sciences, Chair for Formal and Informal Learning, Technical University Munich School of Social Sciences and Technology, Munich, Germany, [2]Aquatic Systems Biology Unit, TUM School of Life Sciences, Technical University of Munich, Freising, Germany

Science education often aims to increase learners' acquisition of fundamental principles, such as learning the basic steps of scientific methods. Worked examples (WE) have proven particularly useful for supporting the development of such cognitive schemas and successive actions in order to avoid using up more cognitive resources than are necessary. Therefore, we investigated the extent to which heuristic WE are beneficial for supporting the acquisition of a basic scientific methodological skill—conducting scientific observation. The current study has a one-factorial, quasi-experimental, comparative research design and was conducted as a field experiment. Sixty two students of a German University learned about scientific observation steps during a course on applying a fluvial audit, in which several sections of a river were classified based on specific morphological characteristics. In the two experimental groups scientific observation was supported either via faded WE or via non-faded WE both presented as short videos. The control group did not receive support via WE. We assessed factual and applied knowledge acquisition regarding scientific observation, motivational aspects and cognitive load. The results suggest that WE promoted knowledge application: Learners from both experimental groups were able to perform the individual steps of scientific observation more accurately. Fading of WE did not show any additional advantage compared to the non-faded version in this regard. Furthermore, the descriptive results reveal higher motivation and reduced extraneous cognitive load within the experimental groups, but none of these differences were statistically significant. Our findings add to existing evidence that WE may be useful to establish scientific competences.

KEYWORDS

digital media, worked examples, scientific observation, motivation, cognitive load

1 Introduction

Learning in science education frequently involves the acquisition of basic principles or generalities, whether of domain-specific topics (e.g., applying a mathematical multiplication rule) or of rather universal scientific methodologies (e.g., performing the steps of scientific observation) (Lunetta et al., 2007). Previous research has shown that worked examples (WE) can be considered particularly useful for developing such cognitive schemata during learning to avoid using more cognitive resources than necessary for learning successive actions (Renkl et al., 2004; Renkl, 2017). WE consist of the presentation of a problem, consecutive solution steps and the solution itself. This is especially advantageous in initial cognitive skill acquisition, i.e., for novice learners with low prior knowledge (Kalyuga et al., 2001). With growing knowledge, fading WE can lead from example-based learning to independent problem-solving (Renkl et al., 2002). Preliminary work has shown the advantage of WE in specific STEM domains like mathematics (Booth et al., 2015; Barbieri et al., 2021), but less studies have investigated their impact on the acquisition of basic scientific competencies that involve heuristic problem-solving processes (scientific argumentation, Schworm and Renkl, 2007; Hefter et al., 2014; Koenen et al., 2017). In the realm of natural sciences, various basic scientific methodologies are employed to acquire knowledge, such as experimentation or scientific observation (Wellnitz and Mayer, 2013). During the pursuit of knowledge through scientific inquiry activities, learners may encounter several challenges and difficulties. Similar to the hurdles faced in experimentation, where understanding the criteria for appropriate experimental design, including the development, measurement, and evaluation of results, is crucial (Sirum and Humburg, 2011; Brownell et al., 2014; Dasgupta et al., 2014; Deane et al., 2014), scientific observation additionally presents its own set of issues. In scientific observation, e.g., the acquisition of new insights may be somewhat incidental due to spontaneous and uncoordinated observations (Jensen, 2014). To address these challenges, it is crucial to provide instructional support, including the use of WE, particularly when observations are carried out in a more self-directed manner.

For this reason, the aim of the present study was to determine the usefulness of digitally presented WE to support the acquisition of a basic scientific methodological skill—conducting scientific observations—using a digital learning environment. In this regard, this study examined the effects of different forms of digitally presented WE (non-faded vs. faded) on students' cognitive and motivational outcomes and compared them to a control group without WE. Furthermore, the combined perspective of factual and applied knowledge, as well as motivational and cognitive aspects, represent further value added to the study.

2 Theoretical background

2.1 Worked examples

WE have been commonly used in the fields of STEM education (science, technology, engineering, and mathematics) (Booth et al., 2015). They consist of a problem statement, the steps to solve the problem, and the solution itself (Atkinson et al., 2000; Renkl et al., 2002; Renkl, 2014). The success of WE can be explained by their impact on cognitive load (CL) during learning, based on assumptions from Cognitive Load Theory (Sweller, 2006).

Learning with WE is considered time-efficient, effective, and superior to problem-based learning (presentation of the problem without demonstration of solution steps) when it comes to knowledge acquisition and transfer (WE-effect, Atkinson et al., 2000; Van Gog et al., 2011). Especially WE can help by reducing the extraneous load (presentation and design of the learning material) and, in turn, can lead to an increase in germane load (effort of the learner to understand the learning material) (Paas et al., 2003; Renkl, 2014). With regard to intrinsic load (difficulty and complexity of the learning material), it is still controversially discussed if it can be altered by instructional design, e.g., WE (Gerjets et al., 2004). WE have a positive effect on learning and knowledge transfer, especially for novices, as the step-by-step presentation of the solution requires less extraneous mental effort compared to problem-based learning (Sweller et al., 1998; Atkinson et al., 2000; Bokosmaty et al., 2015). With growing knowledge, WE can lose their advantages (due to the expertise-reversal effect), and scaffolding learning via faded WE might be more successful for knowledge gain and transfer (Renkl, 2014). Faded WE are similar to complete WE, but fade out solution steps as knowledge and competencies grow. Faded WE enhance near-knowledge transfer and reduce errors compared to non-faded WE (Renkl et al., 2000).

In addition, the reduction of intrinsic and extraneous CL by WE also has an impact on learner motivation, such as interest (Van Gog and Paas, 2006). Um et al. (2012) showed that there is a strong positive correlation between germane CL and the motivational aspects of learning, like satisfaction and emotion. Gupta (2019) mentions a positive correlation between CL and interest. Van Harsel et al. (2019) found that WE positively affect learning motivation, while no such effect was found for problem-solving. Furthermore, learning with WE increases the learners' belief in their competence in completing a task. In addition, fading WE can lead to higher motivation for more experienced learners, while non-faded WE can be particularly motivating for learners without prior knowledge (Paas et al., 2005). In general, fundamental motivational aspects during the learning process, such as situational interest (Lewalter and Knogler, 2014) or motivation-relevant experiences, like basic needs, are influenced by learning environments. At the same time, their use also depends on motivational characteristics of the learning process, such as self-determined motivation (Deci and Ryan, 2012). Therefore, we assume that learning with WE as a relevant component of a learning environment might also influence situational interest and basic needs.

2.1.1 Presentation of worked examples

WE are frequently used in *digital learning scenarios* (Renkl, 2014). When designing WE, the application via digital learning media can be helpful, as their content can be presented in different ways (video, audio, text, and images), tailored to the needs of the learners, so that individual use is possible according to their own prior knowledge or learning pace (Mayer, 2001). Also, digital media can present relevant information in a timely, motivating, appealing and individualized way and support learning in an effective and needs-oriented way (Mayer, 2001). The advantages of using digital media in designing WE have already been shown in previous studies. Dart et al. (2020) presented WE as short videos (WEV). They report that the use of WEV leads to increased student satisfaction and more positive attitudes. Approximately 90% of the students indicated an active learning approach when learning with the WEV. Furthermore, the results show that students improved their content knowledge through WEV and that they found WEV useful for other courses as well.

Another study (Kay and Edwards, 2012) presented WE as video podcasts. Here, the advantages of WE regarding self-determined learning in terms of learning location, learning time, and learning speed were shown. Learning performance improved significantly after use. The step-by-step, easy-to-understand explanations, the diagrams, and the ability to determine the learning pace by oneself were seen as beneficial.

Multimedia WE can also be enhanced with self-explanation prompts (Berthold et al., 2009). Learning from WE with self-explanation prompts was shown to be superior to other learning methods, such as hypertext learning and observational learning.

In addition to presenting WE in different medial ways, WE can also comprise different content domains.

2.1.2 Content and context of worked examples

Regarding the *content* of WE, algorithmic and heuristic WE, as well as single-content and double-content WE, can be distinguished (Reiss et al., 2008; Koenen et al., 2017; Renkl, 2017). Algorithmic WE are traditionally used in the very structured mathematical–physical field. Here, an algorithm with very specific solution steps is to learn, for example, in probability calculation (Koenen et al., 2017). In this study, however, we focus on heuristic double-content WE. Heuristic WE in science education comprise fundamental scientific working methods, e.g., conducting experiments (Koenen et al., 2017). Furthermore, double-content WE contain two learning domains that are relevant for the learning process: (1) the learning domain describes the primarily to be learned abstract process or concept, e.g., scientific methodologies like observation (see section 2.2), while (2) the exemplifying domain consists of the content that is necessary to teach this process or concept, e.g., mapping of river structure (Renkl et al., 2009).

Depending on the WE content to be learned, it may be necessary for learning to take place in different settings. This can be in a formal or informal learning setting or a non-formal field setting. In this study, the focus is on learning scientific observation (learning domain) through river structure mapping (exemplary domain), which takes place with the support of digital media in a formal (university) setting, but in an informal context (nature).

2.2 Scientific observation

Scientific observation is fundamental to all scientific activities and disciplines (Kohlhauf et al., 2011). Scientific observation must be clearly distinguished from everyday observation, where observation is purely a matter of noticing and describing specific characteristics (Chinn and Malhotra, 2001). In contrast to this everyday observation, scientific observation as a method of knowledge acquisition can be described as a rather complex activity, defined as the theory-based, systematic and selective perception of concrete systems and processes without any fundamental manipulation (Wellnitz and Mayer, 2013). Wellnitz and Mayer (2013) described the scientific observation process via six steps: (1) formulation of the research question (s), (2) deduction of the null hypothesis and the alternative hypothesis, (3) planning of the research design, (4) conducting the observation, (5) analyzing the data, and (6) answering the research question(s) on this basis. Only through reliable and qualified observation, valid data can be obtained that provide solid scientific evidence (Wellnitz and Mayer, 2013).

Since observation activities are not trivial and learners often observe without generating new knowledge or connecting their observations to scientific explanations and thoughts, it is important to provide support at the related cognitive level, so that observation activities can be conducted in a structured way according to pre-defined criteria (Ford, 2005; Eberbach and Crowley, 2009). Especially during field-learning experiences, scientific observation is often spontaneous and uncoordinated, whereby random discoveries result in knowledge gain (Jensen, 2014).

To promote successful observing in rather unstructured settings like field trips, instructional support for the observation process seems useful. To guide observation activities, digitally presented WE seem to be an appropriate way to introduce learners to the individual steps of scientific observation using concrete examples.

2.3 Research questions and hypothesis

The present study investigates the effect of digitally presented double-content WE that supports the mapping of a small Bavarian river by demonstrating the steps of scientific observation. In this analysis, we focus on the learning domain of the WE and do not investigate the exemplifying domain in detail. Distinct ways of integrating WE in the digital learning environment (faded WE vs. non-faded WE) are compared with each other and with a control group (no WE). The aim is to examine to what extent differences between those conditions exist with regard to (RQ1) learners' competence acquisition [acquisition of factual knowledge about the scientific observation method (quantitative data) and practical application of the scientific observation method (quantified qualitative data)], (RQ2) learners' motivation (situational interest and basic needs), and (RQ3) CL. It is assumed that (Hypothesis 1), the integration of WE (faded and non-faded) leads to significantly higher competence acquisition (factual and applied knowledge), significantly higher motivation and significantly lower extraneous CL as well as higher germane CL during the learning process compared to a learning environment without WE. No differences between the conditions are expected regarding intrinsic CL. Furthermore, it is assumed (Hypothesis 2) that the integration of faded WE leads to significantly higher competence acquisition, significantly higher motivation, and lower extraneous CL as well as higher germane CL during the learning processes compared to non-faded WE. No differences between the conditions are expected with regard to intrinsic CL.

3 Methods

The study took place during the field trips of a university course on the application of a fluvial audit (FA) using the German working aid for mapping the morphology of rivers and their floodplains (Bayerisches Landesamt für Umwelt, 2019). FA is the leading fluvial geomorphological tool for application to data collection contiguously along all watercourses of interest (Walker et al., 2007). It is widely used because it is a key example of environmental conservation and monitoring that needs to be taught to students of selected study programs; thus, knowing about the most effective ways of learning is of high practical relevance.

3.1 Sample and design

3.1.1 Sample

The study was conducted with 62 science students and doctoral students of a German University (age $M = 24.03$ years; $SD = 4.20$; 36 females; 26 males). A total of 37 participants had already conducted a scientific observation and would rate their knowledge in this regard at a medium level ($M = 3.32$ out of 5; $SD = 0.88$). Seven participants had already conducted an FA and would rate their knowledge in this regard at a medium level ($M = 3.14$ out of 5; $SD = 0.90$). A total of 25 participants had no experience at all. Two participants had to be excluded from the sample afterward because no posttest results were available.

3.1.2 Design

The study has a 1-factorial quasi-experimental comparative research design and is conducted as a field experiment using a pre/posttest design. Participants were randomly assigned to one of three conditions: no WE ($n = 20$), faded WE ($n = 20$), and non-faded WE ($n = 20$).

3.2 Implementation and material

3.2.1 Implementation

The study started with an online kick-off meeting where two lecturers informed all students within an hour about the basics regarding the assessment of the structural integrity of the study river and the course of the field trip days to conduct an FA. Afterward, within 2 weeks, students self-studied via Moodle the FA following the German standard method according to the scoresheets of Bayerisches Landesamt für Umwelt (2019). This independent preparation using the online presented documents was a necessary prerequisite for participation in the field days and was checked in the pre-testing. The preparatory online documents included six short videos and four PDF files on the content, guidance on the German protocol of the FA, general information on river landscapes, information about anthropogenic changes in stream morphology and the scoresheets for applying the FA. In these sheets, the river and its floodplain are subdivided into sections of 100 m in length. Each of these sections is evaluated by assessing 21 habitat factors related to flow characteristics and structural variability. The findings are then transferred into a scoring system for the description of structural integrity from 1 (natural) to 7 (highly modified). Habitat factors have a decisive influence on the living conditions of animals and plants in and around rivers. They included, e.g., variability in water depth, stream width, substratum diversity, or diversity of flow velocities.

3.2.2 Materials

On the field trip days, participants were handed a tablet and a paper-based FA worksheet (last accessed 21st September 2022).[1] This four-page assessment sheet was accompanied by a digital learning environment presented on Moodle that instructed the participants on mapping the water body structure and guided the scientific observation method. All three Moodle courses were identical in structure and design; the only difference was the implementation of the WE. Below, the course without WE are described first. The other two courses have an identical structure, but contain additional WE in the form of learning videos.

1 https://www.lfu.bayern.de/wasser/gewaesserstrukturkartierung/index.htm

3.2.3 No worked example

After a short welcome and introduction to the course navigation, the FA started with the description of a short hypothetical scenario: Participants should take the role of an employee of an urban planning office that assesses the ecomorphological status of a small river near a Bavarian city. The river was divided into five sections that had to be mapped separately. The course was structured accordingly. At the beginning of each section, participants had to formulate and write down a research question, and according to hypotheses regarding the ecomorphological status of the river's section, they had to collect data in this regard via the mapping sheet and then evaluate their data and draw a conclusion. Since this course serves as a control group, no WE videos supporting the scientific observation method were integrated. The layout of the course is structured like a book, where it is not possible to scroll back. This is important insofar as the participants do not have the possibility to revisit information in order to keep the conditions comparable as well as distinguishable.

3.2.4 Non-faded worked example

In the course with no-faded WE, three instructional videos are shown for each of the five sections. In each of the three videos, two steps of the scientific observation method are presented so that, finally, all six steps of scientific observation are demonstrated. The mapping of the first section starts after the general introduction (as described above) with the instruction to work on the first two steps of scientific observation: the formulation of a research question and hypotheses. To support this, a video of about 4 min explains the features of scientific sound research questions and hypotheses. To this aim, a practical example, including explanations and tips, is given regarding the formulation of research questions and hypotheses for this section (e.g., "To what extent does the building development and the closeness of the path to the water body have an influence on the structure of the water body?" Alternative hypothesis: It is assumed that the housing development and the closeness of the path to the water body have a negative influence on the water body structure. Null hypothesis: It is assumed that the housing development and the closeness of the path to the watercourse have no negative influence on the watercourse structure.). Participants should now formulate their own research questions and hypotheses, write them down in a text field at the end of the page, and then skip to the next page. The next two steps of scientific observation, planning and conducting, are explained in a short 4-min video. To this aim, a practical example including explanations and tips is given regarding planning and conducting scientific for this section (e.g., "It's best to go through each evaluation category carefully one by one that way you are sure not to forget anything!"). Now, participants were asked to collect data for the first section using their paper-based FA worksheet. Participants individually surveyed the river and reported their results in the mapping sheet by ticking the respective boxes in it. After collecting this data, they returned to the digital learning environment to learn how to use these data by studying the last two steps of scientific observation, evaluation, and conclusion. The third 4-min video explained how to evaluate and interpret collected data. For this purpose, a practical example with explanations and tips is given regarding evaluating and interpreting data for this section (e.g., "What were the individual points that led to the assessment? Have there been points that were weighted more than others? Remember the introduction video!"). At the end of the page, participants could answer their before-stated research questions and hypotheses by evaluating their collected data and drawing a conclusion. This brings participants to the end of the first mapping section. Afterward, the cycle begins again with the second section of the river that has to be mapped. Again, participants had to conduct the steps of scientific observation, guided by WE videos, explaining the steps in slightly different wording or with different examples. A total of five sections are mapped, in which the structure of the learning environment and the videos follow the same procedure.

3.2.5 Faded worked example

The digital learning environment with the faded WE follow the same structure as the version with the non-faded WE. However, in this version, the information in the WE videos is successively reduced. In the first section, all three videos are identical to the version with the non-faded WE. In the second section, faded content was presented as follows: the tip at the end was omitted in all three videos. In the third section, the tip and the practical example were omitted. In the fourth and fifth sections, no more videos were presented, only the work instructions.

3.3 Procedure

The data collection took place on four continuous days on the university campus, with a maximum group size of 15 participants on each day. The students were randomly assigned to one of the three conditions (no WE vs. faded WE vs. non-faded WE). After a short introduction to the procedure, the participants were handed the paper-based FA worksheet and one tablet per person. Students scanned the QR code on the first page of the worksheet that opened the pretest questionnaire, which took about 20 min to complete. After completing the questionnaire, the group walked for about 15 min to the nearby small river that was to be mapped. Upon arrival, there was first a short introduction to the digital learning environment and a check that the login (via university account on Moodle) worked. During the next 4 h, the participants individually mapped five segments of the river using the cartography worksheet. They were guided through the steps of scientific observation using the digital learning environment on the tablet. The results of their scientific observation were logged within the digital learning environment. At the end of the digital learning environment, participants were directed to the posttest via a link. After completing the test, the tablets and mapping sheets were returned. Overall, the study took about 5 h per group each day.

3.4 Instruments

In the pretest, sociodemographic data (age and gender), the study domain and the number of study semesters were collected. Additionally, the previous scientific observation experience and the estimation of one's own ability in this regard were assessed. For example, it was asked whether scientific observation had already been conducted and, if so, how the abilities were rated on a 5-point scale from very low to very high. Preparation for the FA on the basis of the learning material was assessed: Participants were asked whether they had studied all six videos and all four PDF documents, with the response options not at all, partially, and completely. Furthermore, a factual knowledge test about scientific observation and questions about self-determination theory was administered. The posttest used the same knowledge test, and additional questions on basic needs, situational interest, measures of CL and questions about the usefulness of the WE. All scales were presented online, and participants reached the questionnaire via QR code.

3.4.1 Scientific observation competence acquisition

For the factual knowledge (quantitative assessment of the scientific observation competence), a single-choice knowledge test with 12 questions was developed and used as pre- and posttest with a maximum score of 12 points. It assesses the learners' knowledge of the scientific observation method regarding the steps of scientific observation, e.g., formulating research questions and hypotheses or developing a research design. The questions are based on Wahser (2008, adapted by Koenen, 2014) and adapted to scientific observation: "Although you are sure that you have conducted the scientific observation correctly, an unexpected result turns up. What conclusion can you draw?" Each question has four answer options (one of which is correct) and, in addition, one "I do not know" option.

For the applied knowledge (quantified qualitative assessment of the scientific observation competence), students' scientific observations written in the digital learning environment were analyzed. A coding scheme was used with the following codes: 0 = insufficient (text field is empty or includes only insufficient key points), 1 = sufficient (a research question and no hypotheses or research question and inappropriate hypotheses are stated), 2 = comprehensive (research question and appropriate hypothesis or research question and hypotheses are stated, but, e.g., incorrect null hypothesis), 3 = very comprehensive (correct research question, hypothesis and null hypothesis are stated). One example of a very comprehensive answer regarding the research question and hypothesis is: To what extent does the lack of riparian vegetation have an impact on water body structure? Hypothesis: The lack of shore vegetation has a negative influence on the water body structure. Null hypothesis: The lack of shore vegetation has no influence on the water body structure. Afterward, a sum score was calculated for each participant. Five times, a research question and hypotheses (steps 1 and 2 in the observation process) had to be formulated (5 × max. 3 points = 15 points), and five times, the research questions and hypotheses had to be answered (steps 5 and 6 in the observation process: evaluation and conclusion) (5 × max. 3 points = 15 points). Overall, participants could reach up to 30 points. Since the observation and evaluation criteria in data collection and analysis were strongly predetermined by the scoresheet, steps 3 and 4 of the observation process (planning and conducting) were not included in the analysis.

All 600 cases (60 participants, each 10 responses to code) were coded by the first author. For verification, 240 cases (24 randomly selected participants, eight from each course) were cross-coded by an external coder. In 206 of the coded cases, the raters agreed. The cases in which the raters did not agree were discussed together, and a solution was found. This results in Cohen's $\kappa = 0.858$, indicating a high to very high level of agreement. This indicates that the category system is clearly formulated and that the individual units of analysis could be correctly assigned.

3.4.2 Self-determination index

For the calculation of the self-determination index (SDI-index), Thomas and Müller (2011) scale for self-determination was used in the pretest. The scale consists of four subscales: intrinsic motivation (five items; e.g., I engage with the workshop content because I enjoy it; reliability of alpha = 0.87), identified motivation (four items; e.g., I engage with the workshop content because it gives me more options when choosing a career; alpha = 0.84), introjected motivation (five items; e.g., I engage with the workshop content because otherwise I would have a guilty feeling; alpha = 0.79), and external motivation (three items, e.g., I engage with the workshop content because I simply have to learn it; alpha = 0.74). Participants could indicate their answers on a 5-point Likert scale ranging from 1 = completely disagree to 5 = completely agree. To calculate the SDI-index, the sum of the self-determined regulation styles (intrinsic and identified) is subtracted from the sum of the external regulation styles (introjected and external), where intrinsic and external regulation are scored two times (Thomas and Müller, 2011).

3.4.3 Motivation

Basic needs were measured in the posttest with the scale by Willems and Lewalter (2011). The scale consists of three subscales: perceived competence (four items; e.g., during the workshop, I felt that I could meet the requirements; alpha = 0.90), perceived autonomy (five items; e.g., during the workshop, I felt that I had a lot of freedom; alpha = 0.75), and perceived autonomy regarding personal wishes and goals (APWG) (four items; e.g., during the workshop, I felt that the workshop was how I wish it would be; alpha = 0.93). We added all three subscales to one overall basic needs scale (alpha = 0.90). Participants could indicate their answers on a 5-point Likert scale ranging from 1 = completely disagree to 5 = completely agree.

Situational interest was measured in the posttest with the 12-item scale by Lewalter and Knogler (2014; Knogler et al., 2015; Lewalter, 2020; alpha = 0.84). The scale consists of two subscales: catch (six items; e.g., I found the workshop exciting; alpha = 0.81) and hold (six items; e.g., I would like to learn more about parts of the workshop; alpha = 0.80). Participants could indicate their answers on a 5-point Likert scale ranging from 1 = completely disagree to 5 = completely agree.

3.4.4 Cognitive load

In the posttest, CL was used to examine the mental load during the learning process. The intrinsic CL (three items; e.g., this task was very complex; alpha = 0.70) and extraneous CL (three items; e.g., in this task, it is difficult to identify the most important information; alpha = 0.61) are measured with the scales from Klepsch et al. (2017). The germane CL (two items; e.g., the learning session contained elements that supported me to better understand the learning material; alpha = 0.72) is measured with the scale from Leppink et al. (2013). Participants could indicate their answers on a 5-point Likert scale ranging from 1 = completely disagree to 5 = completely agree.

3.4.5 Attitudes toward worked examples

To measure how effective participants rated the WE, we used two scales related to the WE videos as instructional support. The first scale from Renkl (2001) relates to the usefulness of WE. The scale consists of four items (e.g., the explanations were helpful; alpha = 0.71). Two items were recoded because they were formulated negatively. The second scale is from Wachsmuth (2020) and relates to the participant's evaluation of the WE. The scale consists of nine items (e.g., I always did what was explained in the learning videos; alpha = 0.76). Four items were recoded because they were formulated negatively. Participants could indicate their answers on a 5-point Likert scale ranging from 1 = completely disagree to 5 = completely agree.

3.5 Data analysis

An ANOVA was used to calculate if the variable's prior knowledge and SDI index differed between the three groups. However, as no significant differences between the conditions were found [prior factual knowledge: $F(2, 59) = 0.15$, $p = 0.865$, $\eta^2 = 0.00$ self-determination index: $F(2, 59) = 0.19$, $p = 0.829$, $\eta^2 = 0.00$], they were not included as covariates in subsequent analyses.

Furthermore, a repeated measure, one-way analysis of variance (ANOVA), was conducted to compare the three treatment groups (no WE vs. faded WE vs. non-faded WE) regarding the increase in factual knowledge about the scientific observation method from pretest to posttest.

A MANOVA (multivariate analysis) was calculated with the three groups (no WE vs. non-faded WE vs. faded WE) as a fixed factor and the dependent variables being the practical application of the scientific observation method (first research question), situational interest, basic needs (second research question), and CL (third research question).

Additionally, to determine differences in applied knowledge even among the three groups, Bonferroni-adjusted post-hoc analyses were conducted.

4 Results

The descriptive statistics between the three groups in terms of prior factual knowledge about the scientific observation method and the self-determination index are shown in Table 1. The descriptive statistics revealed only small, non-significant differences between the three groups in terms of factual knowledge.

TABLE 1 Means (standard deviations) of factual knowledge tests (pre- and posttest) and self-determination index for the three different groups.

	No WE	Non-faded WE	Faded WE
Pretest	9.90 (1.17)	10.05 (1.90)	10.10 (1.21)
Posttest	9.65 (1.69)	10.55 (1.23)	10.90 (0.85)
Self-determination index	4.87 (3.20)	5.21 (4.08)	5.52 (2.63)

TABLE 2 Means (standard deviations) of dependent variables with the three different groups.

	No WE	Non-faded WE	Faded WE
Situational interest	3.60 (0.52)	3.76 (0.60)	3.75 (0.72)
Basic needs	3.24 (0.67)	3.58 (0.68)	3.35 (0.77)
Germane load	3.00 (1.01)	3.60 (0.93)	3.60 (0.90)
Intrinsic load	3.10 (0.54)	3.22 (0.84)	3.03 (0.92)
Extraneous load	3.12 (0.71)	2.98 (0.84)	2.55 (0.79)
Applied knowledge (qualitative)	11.65 (4.99)	20.55 (6.89)	19.05 (5.52)

TABLE 3 Means (standard deviations) of dependent variables with the three different groups.

	No WE	Non-faded WE	Faded WE
Usefulness of WE	X	4.10 (0.76)	4.21 (0.68)
Evaluation of WE	X	3.52 (0.74)	3.69 (0.62)

The results of the ANOVA revealed that the overall increase in factual knowledge from pre- to posttest just misses significance [$F(1, 57) = 3.68$, $p = 0.060$, $\eta^2 = 0\ 0.06$]. Furthermore, no significant differences between the groups were found regarding the acquisition of factual knowledge from pre- to posttest [$F(2, 57) = 2.93$, $p = 0.062$, $\eta^2 = 0.09$].

An analysis of the descriptive statistics showed that the largest differences between the groups were found in applied knowledge (qualitative evaluation) and extraneous load (see Table 2).

Results of the MANOVA revealed significant overall differences between the three groups [$F(12, 106) = 2.59$, $p = 0.005$, $\eta^2 = 0.23$]. Significant effects were found for the application of knowledge [$F(2, 57) = 13.26$, $p = <0.001$, $\eta^2 = 0.32$]. Extraneous CL just missed significance [$F(2, 57) = 2.68$, $p = 0.065$, $\eta^2 = 0.09$]. There were no significant effects for situational interest [$F(2, 57) = 0.44$, $p = 0.644$, $\eta^2 = 0.02$], basic needs [$F(2, 57) = 1.22$, $p = 0.302$, $\eta^2 = 0.04$], germane CL [$F(2, 57) = 2.68$, $p = 0.077$, $\eta^2 = 0.09$], and intrinsic CL [$F(2, 57) = 0.28$, $p = 0.757$, $\eta^2 = 0.01$].

Bonferroni-adjusted *post hoc* analysis revealed that the group without WE had significantly lower scores in the evaluation of the applied knowledge than the group with non-faded WE ($p = <0.001$, $M_{diff} = -8.90$, 95% CI [−13.47, −4.33]) and then the group with faded WE ($p = <0.001$, $M_{diff} = -7.40$, 95% CI [−11.97, −2.83]). No difference was found between the groups with faded and non-faded WE ($p = 1.00$, $M_{diff} = -1.50$, 95% CI [−6.07, 3.07]).

The descriptive statistics regarding the perceived usefulness of WE and participants' evaluation of the WE revealed that the group with the faded WE rated usefulness slightly higher than the participants with non-faded WE and also reported a more positive evaluation. However, the results of a MANOVA revealed no significant overall differences [$F(2, 37) = 0.32$, $p = 0.732$, $\eta^2 = 0\ 0.02$] (see Table 3).

5 Discussion

This study investigated the use of WE to support students' acquisition of science observation. Below, the research questions are answered, and the implications and limitations of the study are discussed.

5.1 Results on factual and applied knowledge

In terms of knowledge gain (RQ1), our findings revealed no significant differences in participants' results of the factual knowledge test both across all three groups and specifically between the two experimental groups. These results are in contradiction with related literature where WE had a positive impact on knowledge acquisition (Renkl, 2014) and faded WE are considered to be more effective in knowledge acquisition and transfer, in contrast to non-faded WE (Renkl et al., 2000; Renkl, 2014). A limitation of the study is the fact that the participants already scored very high on the pretest, so participation in the intervention would likely not yield significant knowledge gains due to ceiling effects (Staus et al., 2021). Yet, nearly half of the students reported being novices in the field prior to the study, suggesting that the difficulty of some test items might have been too low. Here, it would be important to revise the factual knowledge test, e.g., the difficulty of the distractors in further study.

Nevertheless, with regard to application knowledge, the results revealed large significant differences: Participants of the two experimental groups performed better in conducting scientific observation steps than participants of the control group. In the experimental groups, the non-faded WE group performed better than the faded WE group. However, the absence of significant differences between the two experimental groups suggests that faded and non-faded WE used as double-content WE are suitable to teach applied knowledge about scientific observation in the learning domain (Koenen, 2014). Furthermore, our results differ from the findings of Renkl et al. (2000), in which the faded version led to the highest knowledge transfer. Despite the fact that the non-faded WE performed best in our study, the faded version of the WE was also appropriate to improve learning, confirming the findings of Renkl (2014) and Hesser and Gregory (2015).

5.2 Results on learners' motivation

Regarding participants' motivation (RQ2; situational interest and basic needs), no significant differences were found across all three groups or between the two experimental groups. However, descriptive results reveal slightly higher motivation in the two experimental groups than in the control group. In this regard, our results confirm existing literature on a descriptive level showing that WE lead to higher learning-relevant motivation (Paas et al., 2005; Van Harsel et al., 2019). Additionally, both experimental groups rated the usefulness of the WE as high and reported a positive evaluation of the WE. Therefore, we assume that even non-faded WE do not lead to over-instruction. Regarding the descriptive tendency, a larger sample might yield significant results and detect even small effects in future investigations. However, because this study also focused on comprehensive qualitative data analysis, it was not possible to evaluate a larger sample in this study.

5.3 Results on cognitive load

Finally, CL did not vary significantly across all three groups (RQ3). However, differences in extraneous CL just slightly missed significance. In descriptive values, the control group reported the highest extrinsic and lowest germane CL. The faded WE group showed the lowest extrinsic CL and a similar germane CL as the non-faded WE group. These results are consistent with Paas et al. (2003) and Renkl (2014), reporting that WE can help to reduce the extraneous CL and, in return, lead to an increase in germane CL. Again, these differences were just above the significance level, and it would be advantageous to retest with a larger sample to detect even small effects.

Taken together, our results only partially confirm H1: the integration of WE (both faded and non-faded WE) led to a higher acquisition of application knowledge than the control group without WE, but higher factual knowledge was not found. Furthermore, higher motivation or different CL was found on a descriptive level only. The control group provided the basis for comparison with the treatment in order to investigate if there is an effect at all and, if so, how large the effect is. This is an important point to assess whether the effort of implementing WE is justified. Additionally, regarding H2, our results reveal no significant differences between the two WE conditions. We assume that the high complexity of the FA could play a role in this regard, which might be hard to handle, especially for beginners, so learners could benefit from support throughout (i.e., non-faded WE).

In addition to the limitations already mentioned, it must be noted that only one exemplary topic was investigated, and the sample only consisted of students. Since only the learning domain of the double-content WE was investigated, the exemplifying domain could also be analyzed, or further variables like motivation could be included in further studies. Furthermore, the influence of learners' prior knowledge on learning with WE could be investigated, as studies have found that WE are particularly beneficial in the initial acquisition of cognitive skills (Kalyuga et al., 2001).

6 Conclusion

Overall, the results of the current study suggest a beneficial role for WE in supporting the application of scientific observation steps. A major implication of these findings is that both faded and non-faded WE should be considered, as no general advantage of faded WE over non-faded WE was found. This information can be used to develop targeted interventions aimed at the support of scientific observation skills.

Data availability statement

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

Ethics statement

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

Author contributions

ML: Writing – original draft. SM: Writing – review & editing. JP: Writing – review & editing. JG: Writing – review & editing. DL: Writing – review & editing.

Funding

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

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2024.1293516/full#supplementary-material

References

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

Barbieri, C. A., Booth, J. L., Begolli, K. N., and McCann, N. (2021). The effect of worked examples on student learning and error anticipation in algebra. *Instr. Sci.* 49, 419–439. doi: 10.1007/s11251-021-09545-6

Bayerisches Landesamt für Umwelt. (2019). *Gewässerstrukturkartierung von Fließgewässern in Bayern – Erläuterungen zur Erfassung und Bewertung. (Water structure mapping of flowing waters in Bavaria - Explanations for recording and assessment)*. Available at: https://www.bestellen.bayern.de/application/eshop_app000005?SID=1020555825&ACTIONxSESSxSHOWPIC(BILDxKEY:%27lfu_was_00152%27,BILDxCLASS:%27Artikel%27,BILDxTYPE:%27PDF%27)

Berthold, K., Eysink, T. H., and Renkl, A. (2009). Assisting self-explanation prompts are more effective than open prompts when learning with multiple representations. *Instr. Sci.* 37, 345–363. doi: 10.1007/s11251-008-9051-z

Bokosmaty, S., Sweller, J., and Kalyuga, S. (2015). Learning geometry problem solving by studying worked examples: effects of learner guidance and expertise. *Am. Educ. Res. J.* 52, 307–333. doi: 10.3102/0002831214549450

Booth, J. L., McGinn, K., Young, L. K., and Barbieri, C. A. (2015). Simple practice doesn't always make perfect. *Policy Insights Behav. Brain Sci.* 2, 24–32. doi: 10.1177/2372732215601691

Brownell, S. E., Wenderoth, M. P., Theobald, R., Okoroafor, N., Koval, M., Freeman, S., et al. (2014). How students think about experimental design: novel conceptions revealed by in-class activities. *Bioscience* 64, 125–137. doi: 10.1093/biosci/bit016

Chinn, C. A., and Malhotra, B. A. (2001). "Epistemologically authentic scientific reasoning" in *Designing for science: implications from everyday, classroom, and professional settings*. eds. K. Crowley, C. D. Schunn and T. Okada (Mahwah, NJ: Lawrence Erlbaum), 351–392.

Dart, S., Pickering, E., and Dawes, L. (2020). Worked example videos for blended learning in undergraduate engineering. *AEE J.* 8, 1–22. doi: 10.18260/3-1-1153-36021

Dasgupta, A., Anderson, T. R., and Pelaez, N. J. (2014). Development and validation of a rubric for diagnosing students' experimental design knowledge and difficulties. *CBE Life Sci. Educ.* 13, 265–284. doi: 10.1187/cbe.13-09-0192

Deane, T., Nomme, K. M., Jeffery, E., Pollock, C. A., and Birol, G. (2014). Development of the biological experimental design concept inventory (BEDCI). *CBE Life Sci. Educ.* 13, 540–551. doi: 10.1187/cbe.13-11-0218

Deci, E. L., and Ryan, R. M. (2012). Self-determination theory. In LangeP. A. M. Van, A. W. Kruglanski and E. T. Higgins (Eds.), *Handbook of theories of social psychology*, 416–436.

Eberbach, C., and Crowley, K. (2009). From everyday to scientific observation: how children learn to observe the Biologist's world. *Rev. Educ. Res.* 79, 39–68. doi: 10.3102/0034654308325899

Ford, D. (2005). The challenges of observing geologically: third graders' descriptions of rock and mineral properties. *Sci. Educ.* 89, 276–295. doi: 10.1002/sce.20049

Gerjets, P., Scheiter, K., and Catrambone, R. (2004). Designing instructional examples to reduce intrinsic cognitive load: molar versus modular presentation of solution procedures. *Instr. Sci.* 32, 33–58. doi: 10.1023/B:TRUC.0000021809.10236.71

Gupta, U. (2019). Interplay of germane load and motivation during math problem solving using worked examples. *Educ. Res. Theory Pract.* 30, 67–71.

Hefter, M. H., Berthold, K., Renkl, A., Riess, W., Schmid, S., and Fries, S. (2014). Effects of a training intervention to foster argumentation skills while processing conflicting scientific positions. *Instr. Sci.* 42, 929–947. doi: 10.1007/s11251-014-9320-y

Hesser, T. L., and Gregory, J. L. (2015). Exploring the Use of Faded Worked Examples as a Problem Solving Approach for Underprepared Students. *High. Educ. Stud.* 5, 36–46.

Jensen, E. (2014). Evaluating children's conservation biology learning at the zoo. *Conserv. Biol.* 28, 1004–1011. doi: 10.1111/cobi.12263

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

Kay, R. H., and Edwards, J. (2012). Examining the use of worked example video podcasts in middle school mathematics classrooms: a formative analysis. *Can. J. Learn. Technol.* 38, 1–20. doi: 10.21432/T2PK5Z

Klepsch, M., Schmitz, F., and Seufert, T. (2017). Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. *Front. Psychol.* 8:1997. doi: 10.3389/fpsyg.2017.01997

Knogler, M., Harackiewicz, J. M., Gegenfurtner, A., and Lewalter, D. (2015). How situational is situational interest? Investigating the longitudinal structure of situational interest. *Contemp. Educ. Psychol.* 43, 39–50. doi: 10.1016/j.cedpsych.2015.08.004

Koenen, J. (2014). *Entwicklung und Evaluation von experimentunterstützten Lösungsbeispielen zur Förderung naturwissenschaftlich experimenteller Arbeitsweisen*. Dissertation.

Koenen, J., Emden, M., and Sumfleth, E. (2017). Naturwissenschaftlich-experimentelles Arbeiten. Potenziale des Lernens mit Lösungsbeispielen und Experimentierboxen. (scientific-experimental work. Potentials of learning with solution examples and experimentation boxes). *Zeitschrift für Didaktik der Naturwissenschaften* 23, 81–98. doi: 10.1007/s40573-017-0056-5

Kohlhauf, L., Rutke, U., and Neuhaus, B. J. (2011). Influence of previous knowledge, language skills and domain-specific interest on observation competency. *J. Sci. Educ. Technol.* 20, 667–678. doi: 10.1007/s10956-011-9322-3

Leppink, J., Paas, F., Van der Vleuten, C. P., Van Gog, T., and Van Merriënboer, J. J. (2013). Development of an instrument for measuring different types of cognitive load. *Behav. Res. Methods* 45, 1058–1072. doi: 10.3758/s13428-013-0334-1

Lewalter, D. (2020). "Schülerlaborbesuche aus motivationaler Sicht unter besonderer Berücksichtigung des Interesses. (Student laboratory visits from a motivational perspective with special attention to interest)" in *Handbuch Forschen im Schülerlabor – theoretische Grundlagen, empirische Forschungsmethoden und aktuelle Anwendungsgebiete*. eds. K. Sommer, J. Wirth and M. Vanderbeke (Münster: Waxmann-Verlag), 62–70.

Lewalter, D., and Knogler, M. (2014). "A questionnaire to assess situational interest – theoretical considerations and findings" in *Poster Presented at the 50th Annual Meeting of the American Educational Research Association (AERA)* (Philadelphia, PA)

Lunetta, V., Hofstein, A., and Clough, M. P. (2007). Learning and teaching in the school science laboratory: an analysis of research, theory, and practice. In N. Lederman and S. Abel (Eds.). *Handbook of research on science education*, Mahwah, NJ: Lawrence Erlbaum, 393–441.

Mayer, R. E. (2001). *Multimedia learning.* Cambridge University Press.

Paas, F., Renkl, A., and Sweller, J. (2003). Cognitive load theory and instructional design: recent developments. *Educ. Psychol.* 38, 1–4. doi: 10.1207/S15326985EP3801_1

Paas, F., Tuovinen, J., van Merriënboer, J. J. G., and Darabi, A. (2005). A motivational perspective on the relation between mental effort and performance: optimizing learner involvement in instruction. *Educ. Technol. Res. Dev.* 53, 25–34. doi: 10.1007/BF02504795

Reiss, K., Heinze, A., Renkl, A., and Groß, C. (2008). Reasoning and proof in geometry: effects of a learning environment based on heuristic worked-out examples. *ZDM Int. J. Math. Educ.* 40, 455–467. doi: 10.1007/s11858-008-0105-0

Renkl, A. (2001). Explorative Analysen zur effektiven Nutzung von instruktionalen Erklärungen beim Lernen aus Lösungsbeispielen. (Exploratory analyses of the effective use of instructional explanations in learning from worked examples). *Unterrichtswissenschaft* 29, 41–63. doi: 10.25656/01:7677

Renkl, A. (2014). "The worked examples principle in multimedia learning" in *Cambridge handbook of multimedia learning*. ed. R. E. Mayer (Cambridge University Press), 391–412.

Renkl, A. (2017). Learning from worked-examples in mathematics: students relate procedures to principles. *ZDM* 49, 571–584. doi: 10.1007/s11858-017-0859-3

Renkl, A., Atkinson, R. K., and Große, C. S. (2004). How fading worked solution steps works. A cognitive load perspective. *Instr. Sci.* 32, 59–82. doi: 10.1023/B:TRUC.0000021815.74806.f6

Renkl, A., Atkinson, R. K., and Maier, U. H. (2000). "From studying examples to solving problems: fading worked-out solution steps helps learning" in *Proceeding of the 22nd Annual Conference of the Cognitive Science Society*. eds. L. Gleitman and A. K. Joshi (Mahwah, NJ: Erlbaum), 393–398.

Renkl, A., Atkinson, R. K., Maier, U. H., and Staley, R. (2002). From example study to problem solving: smooth transitions help learning. *J. Exp. Educ.* 70, 293–315. doi: 10.1080/00220970209599510

Renkl, A., Hilbert, T., and Schworm, S. (2009). Example-based learning in heuristic domains: a cognitive load theory account. *Educ. Psychol. Rev.* 21, 67–78. doi: 10.1007/s10648-008-9093-4

Schworm, S., and Renkl, A. (2007). Learning argumentation skills through the use of prompts for self-explaining examples. *J. Educ. Psychol.* 99, 285–296. doi: 10.1037/0022-0663.99.2.285

Sirum, K., and Humburg, J. (2011). The experimental design ability test (EDAT). *Bioscene* 37, 8–16.

Staus, N. L., O'Connell, K., and Storksdieck, M. (2021). Addressing the ceiling effect when assessing STEM out-of-school time experiences. *Front. Educ.* 6:690431. doi: 10.3389/feduc.2021.690431

Sweller, J. (2006). The worked example effect and human cognition. *Learn. Instr.* 16, 165–169. doi: 10.1016/j.learninstruc.2006.02.005

Sweller, J., Van Merriënboer, J. J. G., and Paas, F. (1998). Cognitive architecture and instructional design. *Educ. Psychol. Rev.* 10, 251–295. doi: 10.1023/A:1022193728205

Thomas, A. E., and Müller, F. H. (2011). "Skalen zur motivationalen Regulation beim Lernen von Schülerinnen und Schülern. Skalen zur akademischen Selbstregulation von Schüler/innen SRQ-A [G] (überarbeitete Fassung)" in *Scales of motivational regulation in student learning. Student academic self-regulation scales SRQ-A [G] (revised version).*

Wissenschaftliche Beiträge aus dem Institut für Unterrichts- und Schulentwicklung Nr. 5 (Klagenfurt: Alpen-Adria-Universität)

Um, E., Plass, J. L., Hayward, E. O., and Homer, B. D. (2012). Emotional design in multimedia learning. *J. Educ. Psychol.* 104, 485–498. doi: 10.1037/a0026609

Van Gog, T., Kester, L., and Paas, F. (2011). Effects of worked examples, example-problem, and problem- example pairs on novices' learning. *Contemp. Educ. Psychol.* 36, 212–218. doi: 10.1016/j.cedpsych.2010.10.004

Van Gog, T., and Paas, G. W. C. (2006). Optimising worked example instruction: different ways to increase germane cognitive load. *Learn. Instr.* 16, 87–91. doi: 10.1016/j.learninstruc.2006.02.004

Van Harsel, M., Hoogerheide, V., Verkoeijen, P., and van Gog, T. (2019). Effects of different sequences of examples and problems on motivation and learning. *Contemp. Educ. Psychol.* 58, 260–275. doi: 10.1002/acp.3649

Wachsmuth, C. (2020). *Computerbasiertes Lernen mit Aufmerksamkeitsdefizit: Unterstützung des selbstregulierten Lernens durch metakognitive prompts. (Computer-based learning with attention deficit: supporting self-regulated learning through metacognitive prompts)*. Chemnitz: Dissertation Technische Universität Chemnitz.

Wahser, I. (2008). *Training von naturwissenschaftlichen Arbeitsweisen zur Unterstützung experimenteller Kleingruppenarbeit im Fach Chemie (Training of scientific working methods to support experimental small group work in chemistry)*. Dissertation

Walker, J., Gibson, J., and Brown, D. (2007). Selecting fluvial geomorphological methods for river management including catchment scale restoration within the environment agency of England and Wales. *Int. J. River Basin Manag.* 5, 131–141. doi: 10.1080/15715124.2007.9635313

Wellnitz, N., and Mayer, J. (2013). Erkenntnismethoden in der Biologie – Entwicklung und evaluation eines Kompetenzmodells. (Methods of knowledge in biology - development and evaluation of a competence model). *Z. Didaktik Naturwissensch.* 19, 315–345.

Willems, A. S., and Lewalter, D. (2011). "Welche Rolle spielt das motivationsrelevante Erleben von Schülern für ihr situationales Interesse im Mathematikunterricht? (What role does students' motivational experience play in their situational interest in mathematics classrooms?). Befunde aus der SIGMA-Studie" in *Erziehungswissenschaftliche Forschung – nachhaltige Bildung. Beiträge zur 5. DGfE-Sektionstagung "Empirische Bildungsforschung"/AEPF-KBBB im Frühjahr 2009*. eds. B. Schwarz, P. Nenninger and R. S. Jäger (Landau: Verlag Empirische Pädagogik), 288–294.