Investigating students' perceived authenticity of learning activities in an out-of-school lab for social sciences: a replication study

Valentina Nachtigall¹ · Nikol Rummel¹

Received: 2 December 2020 / Accepted: 30 July 2021 / Published online: 13 September 2021
© The Author(s) 2021

Abstract

Out-of-school labs (OSLs) aim to foster students' interest in and knowledge about scientific ways of thinking and working by engaging them in authentic activities that emulate processes of scientific inquiry. However, research has not yet focused on investigating whether students perceive those activities as authentic and whether students' perceived authenticity is related to further motivational (e.g. situational interest) or cognitive (e.g. achievement) learning outcomes. An authentic learning activity that emulates scientific inquiry processes is Productive Failure (PF). In PF, students are asked to explore solutions to a complex problem before they have to falsify their solutions during instruction. The present study, which aims to replicate the findings of a previous study, investigates whether PF has an impact on students' perceived authenticity and their situational interest in an OSL for social sciences. We further examined whether students' perceived authenticity is associated with their situational interest and knowledge acquisition. For this purpose, we conducted a quasi-experimental study with 152 10th graders and compared PF to Direct Instruction (DI). DI can be characterized as less authentic, as it promotes the conceptualization of scientific inquiry being nothing else but an application of instructions. The results mostly replicate the findings of the previous study, showing that the authenticity level of the learning activity did neither affect students' perceived authenticity nor their situational interest. Furthermore, students' perceived authenticity correlated with their situational interest but not with their knowledge acquisition. We discuss the results in light of previous research on the effectiveness of authentic learning settings.

Keywords Authentic learning · Out-of-school learning · Perceived authenticity · Situational interest · Productive failure

✉ Valentina Nachtigall
valentina.nachtigall@rub.de; valentina.nachtigall@ruhr-uni-bochum.de
Nikol Rummel
nikol.rummel@rub.de

¹ Educational Psychology Research Group, Institute of Educational Research, Ruhr-Universität Bochum, Universitätsstr. 150, 44801 Bochum, Germany

Introduction

During the last two decades, there have been several attempts to define, conceptualize, and operationalize the buzzword of *authentic learning* (Betz et al., 2016; Fougat et al., 2019; Herrington & Oliver, 2000; Hod & Sagy, 2019; Polman, 2012; Rule, 2006; Shaffer & Resnick, 1999). These attempts suggest several kinds and dimensions of authenticity as well as diverse ways of implementing authenticity within learning contexts. Although authentic learning may be realized in diverse ways, the goal of authentically contextualizing learning experiences often is the same: To demonstrate the real-world relevance and functionality of the targeted learning concepts to learners and to provide them with experiences of how these often abstract concepts are applied in the real out-of-school world (e.g. Brown et al., 1998; Herrington & Oliver, 2000; Lepper, 1988). In other words: Authentic learning aims at overcoming the problem of inert knowledge, meaning that learners often acquire abstract knowledge in school, which they cannot apply in real-life contexts (e.g. Cho et al., 2015; Herrington & Oliver, 2000). By illustrating the relevance and functionality of a targeted learning concept, learners' motivation to engage with the learning content and their understanding of the learning content is expected to be fostered (e.g. Lepper, 1988; Newmann & Wehlage, 1993). These cognitive and motivational benefits of authentically contextualized learning environments are hypothesized to be mediated by learners' perceived authenticity (Betz et al., 2016; Gulikers et al., 2005). Specifically, only if learners perceive a learning environment as authentic, the authentic contextualization affects motivational and cognitive aspects of learning.

Out-of-school learning settings are expected to be highly promising opportunities to offer students authentic learning experiences. Most out-of-school learning settings especially provide authentic contexts for learning and experiencing *science*. Braund and Reiss (2006) as well as Luehmann (2009), for instance, describe various learning settings, such as fieldwork, museums, theme parks, or science enrichment programs, as ideal opportunities to engage learners with authentic practical science, as these settings usually allow learners to conduct hands-on activities with authentic scientific tools. These authentic experiences may be promising for enhancing learners' achievement in and attitudes towards science (Luehmann, 2009). The same applies to out-of-school labs (OSLs), which were numerously initiated in Germany after the *PISA-shock* (Waldow, 2009) in 2001. The PISA results pointed to the rather low scientific literacy of students in Germany. OSLs, which are non-formal learning settings that are usually visited by whole classes during regular school days, try to foster students' scientific literacy by engaging them in authentic experiential learning activities that mimic features and processes of authentic scientific practices. Specifically, OSLs for natural sciences often implement hands-on experimentation activities (e.g. Euler, 2004; Garner & Eilks, 2015) and OSLs for humanities and social sciences usually involve learners in complex and ill-structured problem-based learning activities (Pauly, 2012). Thereby, OSLs aim at triggering students' interest in and enhancing their knowledge about scientific ways of thinking and working (e.g. Scharfenberg & Bogner, 2014).

However, while several studies have been conducted on the motivational effects (e.g. situational interest) and/or cognitive effects (e.g. achievement) of authentic learning settings more generally (e.g. Beier et al., 2019; Smeds et al., 2015; Wilde et al., 2012) and of OSLs more specifically (e.g. Itzek-Greulich et al., 2017; Scharfenberg et al., 2007), a rather small body of research has begun to investigate learners' perceived authenticity (e.g. Betz, 2018; Chang et al., 2010; Gulikers et al., 2005) and thus the role of perceived authenticity for the effects of authentic learning settings, such as OSLs, on motivational

and cognitive learning outcomes. These studies focused on the effect of the authenticity level of the whole learning environment (Gulikers et al., 2005), the learning location (Betz, 2018; Schüttler et al., 2021), or the learning materials (Chang et al., 2010; Sauter et al., 2013; Stamer et al., 2021) on students' perceived authenticity and further motivational (e.g. situational interest) or cognitive (e.g. achievement) learning outcomes. But, so far, it has not to our knowledge been systematically investigated whether students perceive *learning activities* that emulate core features and processes of real-world practices (e.g. scientific practices) as authentic and whether their perceived authenticity affects further outcomes. To be clear, with the term *learning activity* we refer to the method (e.g. problem-based learning, inquiry-based learning, etc.) that students use for approaching a certain learning topic. The fact that the impact of the authenticity level of *learning activities* on students' perceived authenticity and further learning outcomes has not yet been investigated is particularly surprising, as different conceptualizations and operationalizations of authentic learning emphasize the role of learning activities that mirror procedures and processes of real-world practices for creating authentic learning contexts (Betz et al., 2016; Herrington & Oliver, 2000; Rule, 2006). Moreover, as mentioned before, learning activities that enable learners to experience procedures and processes of authentic scientific practices are a major feature of authentic out-of-school learning settings, such as OSLs.

Against this background, which is described in more detail in the following sections, we have begun to investigate whether the authenticity level of a learning activity affects students' perceived authenticity and further outcomes (especially situational interest) while learning in an OSL. More precisely, in the current study, we compared two learning activities: Productive Failure (PF), which is considered a highly authentic learning activity and Direct Instruction (DI), which can be characterized as a less authentic learning activity. Thereby, we aimed at *conceptually replicating* (Schmidt, 2016) the findings of a previous study (Nachtigall et al., 2018). We compared the effects of the two learning activities (i.e., PF versus DI) on students' perceived authenticity and their situational interest while learning the scientific practice of evaluating causal versus correlational evidence in an OSL for social sciences. These analyses are part of a dataset, which is also used in a paper focusing on a different research question and on different variables (Nachtigall et al., 2020).

Theoretical background

Authentic learning in out-of-school labs

Given the theoretical model of authenticity in teaching and learning contexts developed by Betz et al. (2016), OSLs can be characterized as highly authentic learning settings that should—mediated by students' perceived authenticity—provoke several positive effects in students. The model, which is illustrated in simplified form in Fig. 1, is divided into three successive parts. Part 1 highlights the characteristics of learners and of learning settings that may have an impact on students' perceived authenticity of a learning environment. Part 2 outlines the interaction of these two components that results in students' perceived authenticity of the learning environment. Part 3 summarizes the outcomes that may be affected by students' perceived authenticity of the learning environment.

Building on this model, OSLs can be characterized as authentic learning settings, as several features of OSLs contribute to the authentic contextualization of students' learning about scientific ways of thinking and working. Specifically, OSLs are often located

Fig. 1 Model of authenticity in learning contexts adapted from Betz et al. (2016)

at authentic research locations, such as universities, the instructors are often scientists or prospective scientists, and students usually use authentic materials (e.g. original data) and methods (e.g. inquiry-based methods) to work on authentic contents (e.g. current research questions or scientific problems) and open questions (i.e., innovation) (for more detailed characterizations of OSLs, see e.g. Garner & Eilks, 2015; Scharfenberg & Bogner, 2014). Thus, OSLs provide and emulate several features of authentic scientific practices. So far, to the best of our knowledge, eight studies have systematically investigated the effects of features hypothesized to determine the authenticity of an OSL setting on students' motivational or cognitive learning outcomes. Table 1 gives an overview of these studies.

Most of the studies that are listed in Table 1 have examined the effects of the authenticity level of the *location* by comparing student learning in an OSL (i.e., authentic location) to learning in school (non-authentic location) on students' motivational or cognitive learning outcomes. The results with respect to the cognitive effects (i.e., students' knowledge acquisition and achievement) are highly inconsistent. While Scharfenberg et al. (2007) found a positive effect of the OSL setting on students' knowledge acquisition, the findings from Itzek-Greulich et al. (2015) demonstrated no differences between the settings. The results of another study conducted by Itzek-Greulich et al. (2017) even showed negative effects of learning in an OSL compared to learning in school on students' achievement. With regard to the motivational effects of the OSL setting, the findings are again not consistent. While the findings from Betz (2018) demonstrated a positive effect of the location on students' situational interest (mediated by their perceived authenticity), the findings by Itzek-Greulich et al. (2017) and Itzek-Greulich and Vollmer (2017) mainly showed no differences between the learning settings with regard to several motivational outcomes, such as students' situational interest. Also, the study by Schüttler et al. (2021) demonstrated no effect of the location alone on students' situational interest. In their study, the location (i.e., the OSL) provoked only in combination with the provision of authentic learning material (i.e., high-tech lab equipment) positive effects on students' situational interest (Schüttler et al., 2021). Schüttler et al. (2021) further showed that the students who worked with authentic materials reported the highest perceived authenticity, independent from the location. However, OSLs aim to complement and not to replace regular instruction and lessons

Table 1 Overview of OSL-studies that investigated the effects of certain features of an authentic learning setting

#	References	Manipulated authenticity feature	Learning domain	Experimental design	Sample	Results
1	Betz (2018)	Learning location (+ instructor)	Linguistics	Two conditions: Learning in an OSL vs. learning in school	153 8th and 9th graders	Positive effect on students' perceived authenticity (OSL > school) and situational interest (OSL > school)
2	Itzek-Greulich et al. (2015)	Learning location (+ instructor)	Chemistry	Four conditions: Learning in an OSL vs. learning in school vs. learning in both school and OSL (vs. control condition)	1773 9th graders	No effect on students' knowledge acquisition
3	Itzek-Greulich et al. (2017)	Learning location (+ instructor)			1854 9th graders	Negative effect on achievement (OSL < school / OSL and school), no effect on state motivation (e.g. situational interest), but positive effects on some aspects of trait motivation (e.g. competence beliefs: OSL/OSL and school > school)
4	Itzek-Greulich and Vollmer (2017)	Learning location (+ instructor)			1228 9th graders	No effect on motivation (e.g. situational interest) or emotions (e.g. anger)
5	Mierwald et al. (2018)	Learning materials	History science	Three conditions: Learning with original sources vs. learning with an audiobook of the original sources vs. learning with a textbook	161 11th and 12th graders	Positive effect on development of more sophisticated epistemological beliefs (audiobook > original sources > textbook)
6	Scharfenberg et al. (2007)	Learning location and learning activity (method)	Molecular biology	Four conditions: Hands-on experimentation in an OSL vs. no hands-on in an OSL vs. no hands-on in school (vs. control condition)	337 12th graders	No effect of learning activity, but positive effect of location on knowledge acquisition (no hands-on in OSL > no hands-on in school)

Table 1 (continued)

#	References	Manipulated authenticity feature	Learning domain	Experimental design	Sample	Results
7	Schüttler et al. (2021)	Learning location and learning material	Physics	2 × 2 Design: OSL vs. School and high-tech lab equipment vs. simple materials for do-it-yourself experiments	246 10th graders	No effect of location or material on the emotional level and the value-related level of students' situational interest, but positive effect of location and material on the epistemic level of situational interest (OSL + high-tech equipment > all other conditions). Positive effect of material on students' perceived authenticity (high-tech equipment in OSL or school > simple materials in OSL or school)
8	Stamer et al. (2021)	Learning material	Nanotechnology	Two conditions in an OSL: Performing experiments vs. performing experiments and watching videos showing scientists' regular work	236 10th to 13th graders	Positive effect of material (i.e., science videos) on students' beliefs related to scientific practices, and on perceived authenticity

Studies #1, #2, #3, and #4 varied not only the location but at the same time the instructor: While students in school were usually taught by a regular teacher, students in the OSL conditions were taught by a scientist. Studies #2, #3, #4, #6, and #7 implemented a control group without any treatment, which is why we do not describe the results regarding these control groups

in school (Garner & Eilks, 2015). Therefore, one could argue that the question on conditions for successful learning *within* OSLs becomes rather important than the question on the effectiveness of visiting OSLs compared to learning in school more generally.

In addition to the aforementioned study by Schüttler et al. (2021) that not only compared OSL-visits with learning in school but also investigated the role of authentic learning materials for learning *within* OSLs, three further studies have examined conditions for successful learning within OSLs: Scharfenberg et al. (2007), Mierwald et al. (2018), and Stamer et al. (2021). The latter investigated the impact of science videos showing regular practices of scientists on students' perceived authenticity and their beliefs related to the work of scientists. The findings showed that students who performed nanotechnological experiments and watched the science videos reported higher perceived authenticity and developed more adequate beliefs about scientific practices than students who performed the same experiments but without watching the videos. (Stamer et al., 2021). Thus, the authenticity level of the learning materials had an impact on students' perceived authenticity and their beliefs about science in an OSL. Mierwald et al. (2018) examined the effect of authentic learning materials on students' epistemological beliefs about history science. Their results demonstrated that students who learned with original sources or with an audiobook of these original sources developed more sophisticated epistemological beliefs by the end of their OSL visit than students who learned with a rather traditional textbook. Thus, the authenticity level of the learning materials had an impact on students' epistemological beliefs in an OSL. Scharfenberg et al. (2007) compared authentic hands-on experimentation to a (less authentic) learning activity without hands-on experimentation in an OSL for molecular biology. Their results showed no effect of the authenticity level of the learning activity on students' knowledge acquisition.

It can be noted that the eight studies listed in Table 1 differ on various features, such as the investigated learning domains, the assessed outcomes, and the different effects of authenticity on motivational and cognitive outcomes. It can be further noted that only three studies (i.e., Betz, 2018; Schüttler et al., 2021; Stamer et al., 2021) assessed students' perceived authenticity of the learning environment. This is surprising, as the very first investigations related to the effectiveness of OSLs focused on students' perceived authenticity. One of the major findings by Engeln (2004), Glowinski (2007) and Pawek (2009), who evaluated particular OSL-projects for natural sciences, was that students perceived the OSL environment as authentic and that this perceived authenticity significantly predicted or correlated with students' situational interest. A further study, conducted by Damerau (2012), also evaluated different OSL-projects for natural sciences and compared these projects to regular instruction in school (i.e., control condition without treatment). The results demonstrated a large effect of the three OSL-projects (compared to regular instruction in school) on students' perceived authenticity but no effect on their situational interest (Damerau, 2012).

The association between students' perceived authenticity and their situational interest (as demonstrated by Betz, 2018; Engeln, 2004; Glowinski, 2007; Pawek, 2009) is in line with both the authenticity model by Betz et al. (2016) and theories on the development of interest. The authenticity model implies the assumption that an authentically designed learning setting affects students' perceived authenticity and in turn their situational interest or other outcomes (see Fig. 1). Theories on interest development describe situational interest as "a state or an ongoing process during an actual learning activity" (Krapp, 2002, p. 388) and assume that learners' situational interest is first triggered and then maintained by external factors of the learning environment, such as meaningful and relevant learning activities (Hidi & Renninger, 2006). As mentioned above, the authentic contextualization

of learning experiences aims at demonstrating the relevance and functionality of the targeted learning concepts to students. Hence, also in light of theories on the development of interest, it can be expected that an authentically designed and perceived learning setting fosters students' development of situational interest. Furthermore, Lewalter and Geyer (2009) assume that visits of out-of-school settings, such as museums and science centers, promote students' situational interest more generally, and that the authenticity—as one feature of these settings—demonstrates the relevance of the learning contents and thereby particularly evokes students' *maintained* situational interest. However, their findings showed that students' perceived relevance of the contents presented in a museum or science center strongly predicted both their triggered and maintained situational interest (Lewalter & Geyer, 2009). Against this background, it is reasonable to assume that students' perceived authenticity of an OSL is associated with their (triggered and maintained) situational interest.

However, it is further surprising that, to our knowledge, only Scharfenberg et al. (2007) have investigated the effectiveness of an authentic learning activity (i.e., authentic hands-on experimentation versus an activity without authentic hands-on experimentation) for learning in an OSL. This is surprising, because the involvement of students in authentic learning activities that mimic scientific inquiry practices is a fundamental characteristic of OSLs (e.g. Euler, 2004; Pauly, 2012; Scharfenberg & Bogner, 2014). Engaging learners in inquiry-based learning activities is often also part of characterizations of authentic learning environments more generally (see Betz et al., 2016; Herrington & Oliver, 2000; Rule, 2006). But it has not yet been investigated whether students perceive activities that are assumed to simulate scientific inquiry processes as authentic and whether their perceived authenticity affects further outcomes, such as their situational interest. The present study aims to address this research gap by investigating the role of the authenticity level of a learning activity for students' perceived authenticity and their situational interest in an OSL.

Authentic learning activities

The learning activities implemented in OSLs focus on “getting school students an authentic feeling for the scientific endeavor” (Euler, 2004, p. 26) and try to “place students in the role of being a scientist” (Glowinski & Bayrhuber, 2011, p. 374). The opportunities to situate learners in the role of a researcher by implementing respective learning activities differ depending on the discipline. OSLs for natural sciences focus on implementing hands-on activities that enable students to experiment with authentic tools and materials (e.g. Euler, 2004; Glowinski & Bayrhuber, 2011). OSLs for social sciences focus on implementing problem-based learning activities, which ask students to work independently on real-world and open-ended problems, in order to enable them to undertake research methods in these disciplines (Pauly, 2012).

In order to involve students in authentic activities that simulate features and processes of scientific inquiry, the PF (Productive Failure) activity seems particularly promising. PF asks students to develop conjectures about a complex and novel problem during an initial problem-solving phase before they receive instruction on canonical solutions (e.g. Kapur, 2015). Thus, PF students usually fail and struggle to canonically solve a complex problem. These failure experiences may be accompanied by uncertainty and frustration. Afterwards (i.e., during instruction), however, it is assumed that students discover the limitations of their previously posed conjectures and in turn learn from their failure (see Loibl et al.,

2017). These processes demonstrate that PF students probably approach their understanding about a targeted learning concept just in the same way as scientists usually approach their knowledge about a given problem: Through failure and by discovering the boundaries of their current knowledge (see Firestein, 2015) and by falsifying their previously posed conjectures (see Chalmers, 2013). In addition, it is argued that PF emulates the scientific inquiry process of exploring solution ideas to a complex and novel problem without knowing whether one's conjectures about potential solutions will be supported or not (see Cho et al., 2015). By emulating these major features and processes of scientific inquiry, PF holds the potential to situate learners in the role of a scientist (see Kapur & Toh, 2015) and can be characterized as an authentic learning activity. In previous studies, PF has typically been compared to DI (for a review, see Loibl et al., 2017). In contrast to PF, the DI activity can be characterized as less authentic. Due to students' involvement in instruction first, followed by problem solving, DI approaches promote the conceptualization of scientific inquiry being nothing else but an application of instructions and thus a "simple, algorithm procedure" (Hodson, 1999, p. 784).

Besides emulating core features and processes of scientific inquiry, PF also fulfills several features that characterize authentic learning activities more generally. That is, several operationalizations of authentic learning activities (Betz et al., 2016; Herrington & Oliver, 2000; Rule, 2006) describe that authentic activities should involve complex problem-solving tasks that have to be independently and collaboratively investigated by learners through open inquiry. These features of authentic learning activities (i.e., complex problem, inquiry-based learning activity, collaboration between learners, and independent working) put forward by different theoretical models are in line with the design principles of PF (see Table 2).

As Kapur and Bielaczyc (2012) argue, the initial problem-solving phase of PF should match the following design principles: (1) the PF problem should be complex and open to various solutions (design principle: Problem/task), and (2) students should be able to explore and generate intuitive solution ideas to this problem by using their prior knowledge (design principle: Activity). (3) Students' exploration activities during the problem-solving phase should be fostered by small-group collaboration (design principle: Participation structure) without instructional support from teachers or other facilitators, so that (4) students can independently explore solution ideas (design principle: Social surround). While PF matches several features of authentic learning activities, the problem-solving phase of DI fulfills none of these features (see Table 2). That is, the problem-solving task lacks complexity and novelty, as students usually only have to work on problems that are identical or isomorphic to the problem introduced during instruction (design feature: Problem/task). Due to the initial instruction phase, DI students also do not need to explore solution ideas, but rather have to apply and execute the problem-solving steps they were instructed on (design feature: Activity). Thus, they do not need to engage in a collaborative discourse about their shared solution ideas (design feature: Participation structure), and they do not have to independently develop and choose own problem-solving paths (design feature: Social surround). Therefore, also in light of common features of authentic learning activities, the PF learning activity seems to be more authentic than the learning activity of DI.

The previous study

As outlined before, numerous definitions and characterizations of OSLs and authentic learning settings more generally emphasize the role of certain learning activities in order to

Table 2 Comparison of design features between PF, authentic learning activities, and DI

	Design principles of the PF problem-solving phase	Features of authentic learning activities	Characteristics of the DI problem-solving phase
Problem/task	<i>Complex and novel problems</i> that are open to various and multiple solutions and which address students' prior knowledge	<i>Complex and ill-defined tasks</i> that address real-world problems and questions	Problems that are <i>identical and/or isomorphic</i> to problems previously introduced during instruction
Activity	<i>Exploration</i> and generation of intuitive solution ideas prior to instruction	Simulation of professional practices (e.g. research processes) by engaging students in <i>open inquiry</i>	<i>Application</i> of previously instructed steps for solving a problem
Participation structure	<i>Small-group collaboration</i> that engages students in discourse about their shared solution ideas	Students <i>collaboratively</i> unravel a problem and engage in a discourse	<i>Collaborative performance</i> of previously instructed problem-solving steps
Social surround	<i>Safe place to independently explore</i> various intuitive solution ideas without instructional support	<i>Student-centered</i> activities that empower students to independently define and choose the problem and its steps of solution	Students <i>independently exercise</i> problem-solving paths they were instructed on

design for authenticity. Authentically designed learning settings are hypothesized to affect students' perceived authenticity and, in turn, different motivational and cognitive learning outcomes (Betz et al., 2016; Gulikers et al., 2005). Research, however, has not yet focused on investigating whether students indeed perceive learning activities that are intended to be authentic as authentic and whether their perceived authenticity is related to further learning outcomes, such as students' situational interest and knowledge acquisition. Against this background, we investigated the role of the authenticity level of learning activities in fostering students' perceived authenticity and situational interest.

For this purpose, we compared the effects of PF as an authentic learning activity to DI as a less authentic activity on students' perceived authenticity and their situational interest while learning the scientific practice of designing experiments in an OSL for social sciences (Nachtigall et al., 2018). Our results demonstrated that PF students did not report higher perceived authenticity nor higher situational interest than DI students. In addition, our findings showed that students' perceived authenticity correlated with their situational interest but not with their knowledge acquisition in both conditions.

Based on these findings, we hypothesized that the authenticity level of other features of the learning setting (e.g. university campus, prospective scientist as an instructor, etc.), which were the same in both conditions, might have had a higher impact on students' perceived authenticity than the authenticity level of the learning activity (i.e., PF versus DI). Specifically, due to a potential interrelatedness of the different features of an authentic learning setting, other authentic elements of the OSL (which were the same in both conditions) might have outweighed the effect of the authenticity level of the learning activity on students' perceived authenticity and further outcomes.

Moreover, we assumed that school students may lack knowledge about research methods in the social sciences and especially about the notion that failure, struggle, mistakes, and uncertainty are an inherent part of scientific inquiry. Consequently, students may have difficulties to perceive the PF activity as more authentic than DI.

The present study

To address the aforementioned issues, we conducted a second study and implemented a slightly different design. We again compared the effects of PF with DI on students' perceived authenticity and their situational interest while learning scientific practices in an OSL for social sciences. However, in the present study, we introduced various research methods within the social sciences, briefly described the steps of a typical research process, and communicated the notion that failure and mistakes are an inherent part of scientific inquiry to students in both conditions. Thereby, we aimed at addressing students' potentially limited knowledge about social science research methods. We, moreover, selected a different learning topic: The scientific practice of evaluating causal versus correlational evidence. This topic was particularly appropriate in order to exemplify that researchers also make mistakes, such as interpreting correlational evidence as causal. To teach students the concepts of causality and correlation, the learning materials referred to a real study (i.e., a study that is published in a scientific journal) on the relation between children's usage of media with violent content and their aggressive behavior. This content allowed for demonstrating that research can result in inconsistent, contradicting, and unclear evidence.

By using this slightly modified design in the present study, we aimed at replicating the findings of our first study. Specifically, we pursued what Schmidt (2016) calls a *conceptual replication*. Other than in a direct replication, we did not repeat the same procedure of our

previous study but tested whether or not different learning materials and a slightly different procedure would lead to a repetition of our previous findings.

Research questions and hypotheses

As in our previous study, our major research question (RQ1) is whether the authenticity level of the learning activity (PF vs. DI) has an impact on students' perceived authenticity and whether their perceived authenticity is associated with further outcomes, namely students' situational interest and their performance on a knowledge test. Specifically, with respect to our RQ1, we derived and investigated the following four hypotheses:

Hypothesis 1 (H1) PF students will report higher perceived authenticity at the end of their OSL visit than DI students. H1 builds on the argument that PF emulates features of authentic scientific inquiry and fulfills more features of authentic learning activities than DI (see Table 2).

Hypothesis 2 (H2) PF students will report higher situational interest than DI students at the end of their OSL visit. H2 is based on theories on both authentic learning (e.g. Betz et al., 2016; Lepper, 1988) and interest development (e.g. Hidi & Renninger, 2006), which suggest that the authenticity level of the learning activity has an effect on students' situational interest.

Hypothesis 3 (H3) Students' perceived authenticity will positively correlate with their situational interest, or the effect of the authenticity level of the learning activity on students' situational interest will be even mediated by their perceived authenticity. H3 relies on the findings of previous OSL research (e.g. Betz, 2018) showing that students' perceived authenticity correlates or predicts their situational interest.

Hypothesis 4 (H4) Students' perceived authenticity will positively correlate with their learning outcome (i.e., their performance on a knowledge test). H4 builds on the theoretical assumption that the authentic contextualization of learning fosters students' understanding (e.g. Betz et al., 2016; Lepper, 1988). However, it should be mentioned that PF students in the present study did not outperform DI students on a knowledge test (see our parallel study on the same dataset: Nachtigall et al., 2020). However, the findings of our previous study demonstrated no correlation between students' perceived authenticity and their performance on a knowledge test. Thus, it is interesting to examine whether the theoretically assumed association between students' perceived authenticity and their knowledge acquisition can be empirically supported, and whether or not the present study replicates the findings of our previous study.

In addition to testing these four hypotheses, we aim to explore whether students perceived authenticity of the learning activity (i.e., PF versus DI) is related to their perceived authenticity of other features of the learning setting (RQ2). For this purpose, we administered an additional instrument in the present study (which was not used in our previous study) for assessing students' perceived authenticity of different features of the learning setting.

Materials and methods

Participants

Participants were 10th graders from six secondary schools. 152 students (Age: $M = 16.11$ $SD = 0.90$; 65% girls) from seven social or educational science classes agreed to participate with written parental consent. The classes were randomly assigned to the PF and DI condition as a whole, resulting in the following subsamples: Three classes with a total sample of 80 students in the PF condition and four classes with a total sample of 72 students in the DI condition.

Quasi-experimental design and procedure

The study took place in an interdisciplinary OSL at a large German university which is usually visited by whole classes, leading to the quasi-experimental design of our study (i.e., assignment of whole classes to the conditions).

On the day of the intervention, students in both conditions received an introduction on different research methods within the social sciences and the steps of a typical research process in the social sciences. Afterwards, they experienced two successive learning phases: A problem-solving phase and an instruction phase. In the PF condition, students received instruction after problem solving. In the DI condition, students received instruction prior to problem solving. Thus, the sequence of the two learning phases differed between the two conditions. At the end of the second learning phase, however, students in both conditions experienced a practice phase.

During problem-solving, students in both conditions were asked to form small groups of three or in exceptional cases, due to class size constraints, of two or four. Students usually collaborated with their seat neighbors. The problem-solving task asked them to think about possible results and interpretations of a correlation study. For this purpose, they received a description of the design of a real study (i.e., published in a scientific journal) on the relation between exposure to media violence and aggressive behavior. The problem (violence-in-media problem) was the same in both conditions. Thus, while PF students had to develop intuitive solution ideas, DI students were asked to apply the previous instructions. Students in both conditions did not receive any instructional support during problem solving.

During instruction, students in both conditions received information on the different results that can be retrieved from a correlation study (i.e., positive, negative, and no correlation), a description of correct if-then interpretations of correlational evidence, and an explanation why correlational evidence cannot be interpreted as causal. Moreover, it was emphasized that a common mistake of students (i.e., to evaluate correlational evidence as causal evidence) is widely spread in research, and that failure and mistakes are an inherent part of research more generally. Finally, the instructor presented a potential experimental study design which could offer causal evidence under certain circumstances. All instructional explanations and information were the same in both conditions and embedded in the theme of the violence-in-media problem. The only difference between the two conditions was that the violence-in-media problem was introduced in detail in the DI condition, but only briefly recapitulated in the PF condition.

During the practice phase, students were asked to individually solve a problem isomorphic to the violence-in-media problem. Afterwards, the instructor briefly presented the canonical solution.

Prior to the first learning phase and after both learning phases, we administered online questionnaires. At the end of the OSL-visit, we administered an online knowledge test. Students filled in the questionnaires and worked on the knowledge test on laptops. Table 3 provides an overview of the quasi-experimental design and procedure.

Learning materials

To teach the scientific practice of evaluating causal versus correlational evidence, we designed an instructional lesson and two problem-solving tasks: A problem that students were asked to solve during the collaborative problem-solving phase and an isomorphic problem for the practice phase that students had to solve individually. The two problem-solving tasks each introduced the design of a published correlation study. The problem provided during the problem-solving phase referred to a published study on the relation between exposure to violent media and aggressive behavior. The problem that students were asked to solve individually during the practice phase referred to a published study on the relation between primary school students' usage of the formal (in German: *Sie*) versus informal (in German: *Du*) form to address teachers and their achievement. In Germany, the formal *Sie* is usually used to address strangers, people in authority, etc. and the informal *Du* is used when speaking to friends, family members, and people who offered the *Du*. In both learning phases, the problem-solving tasks asked students to think about potential results of the study and respective interpretations. Moreover, they were asked to discuss whether or not the study design allowed for testing a causal hypothesis. The instructional lesson addressed the correct interpretation of different correlation study results, the issue why correlational evidence cannot be interpreted as causal evidence, and how a study has to be designed such that it can reveal causal evidence. The learning materials were the same in both conditions.

Variation of the authenticity level

Given the characteristics of an authentic learning setting listed in the authenticity model by Betz et al. (2016), we varied—by comparing two learning activities (i.e., PF and DI)—the authenticity level of the method that students used to work on a problem-solving task. The other features of the learning setting (i.e., location, instructor, content, materials, and innovation) were the same in both conditions. Table 4 provides an overview of how the features hypothesized to determine the authenticity of a learning setting are implemented in both conditions.

Measures

A pre-questionnaire (administered prior to the first learning phase) assessed students' individual interest in the subject of social sciences as a control variable. Students' individual interest in a topic or subject is defined "as a relatively stable tendency to occupy oneself with an object of interest" (Krapp, 2002, p. 388) and is assumed to affect students' situational interest. To assess students' individual interest in social sciences, we adapted seven

Table 3 Overview of the quasi-experimental design and procedure

	PF	DI
Introduction	Description of different research methods and the steps of a typical research process within the social sciences	
Pre-questionnaire	Assessed student interest in social sciences (and further variables that are not focused on in this paper)	
Learning phase 1	Problem solving in small groups (violence-in-media problem)	Instruction on canonical solution
Process-questionnaire	Assessed variables that are not focused on in this paper	
Learning phase 2	Instruction on canonical solution	Problem solving in small groups (violence-in-media problem)
(+ practice phase)	Individual practicing (+ brief instructional explanation)	
Post-questionnaire	Assessed perceived authenticity and situational interest (and further variables that are not focused on in this paper)	
Knowledge test	Assessed knowledge about evaluating causal versus correlational evidence	

Table 4 Authenticity features implemented in both conditions

Feature of the setting	PF	DI
Method (learning activity)	Collaborative problem solving prior to instruction: Students have to explore and develop solution ideas to a complex and novel problem without knowing whether or not their conjectures will be supported	Instruction followed by collaborative problem solving: Students have to solve the same problem that was introduced during instruction by applying and executing the instructional explanations
Location	Out-of-school lab on a university campus	
Instructor	Educational research scientist	
Content	Difference between causal versus correlational evidence (with regard to research on whether children's usage of media with violent content is related to/causes aggressive behavior)	
Materials	Laptops, descriptions of real (i.e., published) studies (in two problem-solving tasks), and instructional lesson	
Innovation	During instruction, the instructor emphasized that there has not yet been conclusive evidence on the relation between exposure to media violence and aggressive behavior. Thus, during the problem-solving phase, students discussed a research question that has not yet been answered	

items from a scale developed by Sparfeldt et al. (2004). Specifically, the questionnaire by Sparfeldt et al. (2004) measures students' interest in four subjects (i.e., Mathematics, Physics, German language, and English language) with eight items. In a later study, Rost et al. (2008) used only seven items of this scale for measuring students' subject-specific interest. We adapted these seven items for the social sciences context.

A post-questionnaire (administered after the second learning phase) assessed students' perceived authenticity of the learning activity, their situational interest, and their perceived authenticity of different features of the learning setting.

To assess students' *perceived authenticity of the learning activity*, we adapted ten items from a questionnaire developed by Gulikers et al. (2006). Specifically, Gulikers et al. (2006) developed a questionnaire with originally eight and after validation six scales for assessing students' perceptions of an authentic assessment. Three scales relate to the perception of the assessment criteria and, thus, did not fit our purpose of measuring students' perceived authenticity of the learning activity. The remaining three scales relate to the perceived authenticity of the form (e.g. "this way of assessing fits well with the social work profession), the physical context (e.g. "the context in which I had to perform this assessment looked like the professional practice of a social worker"), and the task (e.g. "the task of this assessment resembled the tasks of a real social worker") of the assessment. These scales (with, in total, 14 items) seemed transferable to our context and purpose. Therefore, we changed the wording of the items and used the term *task* or *project* instead of the term *assessment*. Instead of the social work context, our items referred to the profession of a scientist. After this adaption, we excluded four items that appeared redundant due to their high similarity with other items. A principal axis factor analysis with oblique rotation (direct oblimin) revealed two factors with an eigenvalue of 1 or more, accounting for 48.65% of the variance. Most of the items (i.e., eight items) loaded on one factor, while the second factor united only two items. As in our previous study (see Nachtigall et al., 2018), we, therefore, excluded these two items and formed one scale with eight items for our analysis of students' perceived authenticity of the learning activity.

To measure situational interest, we adapted the scale from Lewalter and Willems (2009). For a full list of nearly similar items, see Knogler et al. (2015). The scale distinguishes between a catch and a hold dimension of situational interest, which is based on the assumptions by Hidi and Renninger (2006) that situational interest is first triggered (catch dimension) and then maintained (hold dimension). The scale from Lewalter and Willems (2009) uses six (catch) items to assess triggered situational interest and six (hold) items to measure maintained situational interest. As Lewalter and Willems (2009) administered their questionnaire in a mathematics classroom, we slightly adapted the wording of the items for the learning environment of our study.

To measure students' perceived authenticity of different features of the learning setting, we used an instrument developed by Wirth et al. (2017). Coming from the authenticity model put forward by Betz et al. (2016), the instrument assesses students' perceived authenticity of the location (two items), the instructor (three items), the material (three items), the method (three items), the content (two items), and the innovation (one item). Moreover, three items assess students' experienced curiosity and excitement while being involved in scientific inquiry activities.

Students replied on a scale of 1 (strongly disagree) to 5 (strongly agree) to each item of the four instruments. Specifically, for reasons of standardization, we employed a five-point Likert scale in all questionnaires, whereby our items assessing students' individual interest differ from the six-point Likert scale used by Sparfeldt et al. (2004). The other three instruments have also been used with five-point Likert scales in previous research (see Gulikers

et al., 2008; Lewalter & Willems, 2009; Wirth et al., 2017). Example items and the internal consistencies of all scales are depicted in Table 5. As shown in Table 5, the instruments that assessed students' individual interest, situational interest (catch as well as hold), and perceived authenticity of the learning activity reach acceptable levels of internal consistencies, as Cronbach's alphas are higher than .70 (Cicchetti, 1994). The scales of the questionnaire developed by Wirth et al. (2017) reach less acceptable internal consistencies between .63 and .76.

To assess students' knowledge about causal versus correlational evidence, we administered a self-developed knowledge test after the second learning phase and a one-hour lunch break. The test was the same in both conditions, contained 10 items, and students worked individually on the test items in a prescribed order (on a laptop). The items of the knowledge test asked students to reproduce, apply, or transfer the concepts they were instructed on during the instruction phase. Most of the items (i.e., 8 out of 10) required students to use their own words in order to, for instance, name possible results that could be revealed by a correlation study, to interpret different correlational results of a fictitious study, or to reason why correlational evidence cannot be interpreted as causal. Two raters coded around 20% of the knowledge tests (i.e., $n=35$) and reached a high agreement between their ratings ($ICC_{absolute} = .93$; 95%-CI [.85, .96]). Prior to our analyses, we excluded two items due to negative correlations with other items of the knowledge test. The total score of the remaining eight items ranged from 0 to 21. Table 6 gives a description of the final eight items of the posttest. For a detailed description of the posttest items, see the supplementary materials in Nachtigall et al. (2020).

Results

Research question 1

To examine the effect of the authenticity level of the learning activity (PF vs. DI) on students' perceived authenticity and situational interest and to investigate the relation between students' perceived authenticity and further learning outcomes (RQ1), we conducted variance and mediation analyses and calculated correlations. To account for a potential alpha-error inflation, we tested the hypotheses that related to students' perceived authenticity with an adjusted significance level of $\alpha=.013$. Specifically, we divided $\alpha=.05$ by the four¹ hypotheses that referred to students' perceived authenticity (see Chen et al., 2017).

Prior to our analyses, we examined whether our data fulfilled crucial prerequisites of parametric analyses. Firstly, we conducted Shapiro–Wilk tests in order to test for normal distribution of our data. The results revealed that the distribution for all four dependent variables appeared to be normal (i.e., $p>.05$) in both conditions. Specifically, for perceived authenticity, the PF data, $D(80)=0.98$, $p=.18$, and the DI data, $D(72)=0.97$, $p=.12$, were

¹ As we used two scales of situational interest, we tested the following four hypotheses with regard to students' perceived authenticity: (1) PF students report higher perceived authenticity than DI students, (2) students' perceived authenticity positively correlates with their triggered situational interest, (3) students' perceived authenticity positively correlates with their maintained situational interest, (4) students' perceived authenticity positively correlates with their performance on a knowledge test. Our mediation analyses serve as an alternative summary analysis of the aforementioned hypotheses. Thus, we did not account for these tests when adjusting the alpha level.

Table 5 Example items and internal consistencies of the questionnaire scales

Scale	Adapted from...	Example item	Cronbach's α
Individual interest	Sparfeldt et al. (2004)/ Rost et al. (2008)	The subject of social sciences is one of my favorite subjects	.87
Situational interest	Lewalter and Willems (2009)/Knogler et al. (2015)	To what extent was the engagement with the contents of this project exciting for you? To what extent would you like to learn more about certain topics of the project?	.88 .77
Perceived authenticity			
Learning activity	Gulikers et al. (2006)	I think that the tasks in this project fit well with the work of real scientists	.86
Location	Wirth et al. (2017)	I worked at a location where research takes place	.70
Instructor		My instructor was a real scientist	.76
Method		I exactly worked as one does in research	.64
Content		I worked on a scientific topic	.68
Materials		I worked with things that are also used in research	.63
Innovation		I worked on a question to which research has not yet an answer	(Only one item)
Excitement and curiosity		It was exciting to me not to know whether my conjectures would be supported or not in the end	.66

Table 6 Description of the final eight posttest item (see also Nachtigall et al., 2020)

Description	Points
Item #1.1 asked students to name the results that could be revealed from a fictitious correlation study (i.e., positive, negative, and no correlation between variables x and y)	3
Item #1.2 asked students to interpret the hypothetical results named in respond to item #1.1 (e.g. If variable x is high, then variable y is high)	3
Item #1.3 asked students to explain whether and why or why not one of the results (named in respond to the previous items) could confirm a causal relation between two variables	2
Item #2.1 presented students with two correct and three incorrect (i.e., causal) interpretations of a positive correlation between two variables and asked them to mark the correct interpretations	2
Item #2.2 asked students to state and reason in their own words how a positive correlation between two variables could <i>not</i> be interpreted (i.e., not as causal evidence due to confounding factors)	2
Item #2.3 required students to describe what kind of study researchers would have to conduct in order to test a causal relation between two variables	5
Item #4 and item #5 each described a fictitious experimental study design and asked the students to identify two problems within each of the two designs	4 (two points per item)

normal. Also, for situational interest (catch), the PF data, $D(80)=0.98$, $p=.25$, and DI data, $D(72)=0.98$, $p=.35$, were normally distributed. For situational interest (hold), the PF data, $D(80)=0.97$, $p=.06$, as well as the DI data, $D(72)=0.98$, $p=.37$, again were normal. Finally, for knowledge-test performance, the PF data, $D(80)=0.98$, $p=.34$, and the DI data, $D(72)=0.99$, $p=.53$, also appeared to be normal. Secondly, we conducted a Levene's test in order to examine whether our data for perceived authenticity and situational interest met the assumption of homogeneity of variance (which is a requirement for conducting variance analyses). For perceived authenticity, the results revealed equal variances for PF and DI, $F(1, 150)=0.10$, $p=.75$. For situational interest (catch), the Levene's test again showed equal variances for PF and DI, $F(1, 150)=0.01$, $p=.94$. Also for situational interest (hold), the results confirmed homogeneity of variance between PF and DI, $F(1, 150)=3.40$, $p=0.07$. In summary, our data met crucial requirements for conducting variance and correlation analyses.

Group differences

To test our H1 (i.e., PF students report higher perceived authenticity than DI students) and our H2 (i.e., PF students report higher situational interest than DI students), we conducted a MANCOVA with the factor condition and students' perceived authenticity of the learning activity and the two scales of situational interest as dependent variables. We included students' interest in the social sciences as a covariate due to correlations with situational interest (catch: $r=.27$, $p=.001$; hold: $r=.26$, $p=.002$). The findings revealed no significant effects of condition on neither of our dependent variables. Thus, against our H1 and H2, but in line with the findings of our first study, the authenticity level of the learning activity did not affect students' perceived authenticity nor their situational interest. Table 7 provides the descriptive statistics, the results of the MANCOVA, and the confidence intervals of the effect sizes, which we calculated for a better interpretation of the null effects (Aberson, 2002).

Table 7 Overview of descriptive statistics and the results of the MANCOVA regarding PF and DI students' perceived authenticity and situational interest

	PF (<i>n</i> = 80) <i>M</i> (<i>SD</i>)	DI (<i>n</i> = 72) <i>M</i> (<i>SD</i>)	<i>F</i> [1, 149]	<i>p</i>	η_p^2	90% CI
Perceived authenticity of the learning activity	3.13 (0.62)	3.31 (0.70)	2.90	.09	.019	[-.00; .07]
Situational interest (catch)	3.14 (0.69)	3.16 (0.72)	0.06	.80	.000	[-.00; .02]
Situational interest (hold)	2.89 (0.65)	2.85 (0.76)	0.11	.75	.001	[-.00; .02]

Table 8 Correlations between students' perceived authenticity of the learning activity, situational interest, and knowledge-test performance

Pair of variables	All (<i>N</i> = 152)		PF (<i>n</i> = 80)		DI (<i>n</i> = 72)	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Perceived authenticity—situational interest (catch)	.30	<.001	.17	.14	.42	<.001
Perceived authenticity—situational interest (hold)	.38	<.001	.22	.05	.52	<.001
Perceived authenticity—knowledge-test performance	.14	.08	.28	.012	.00	.99

Correlations

To test our H3 (i.e., students' perceived authenticity is associated with their situational interest) and our H4 (i.e., students' perceived authenticity is associated with their performance on the knowledge test²), we conducted correlation analyses. Table 8 provides the results of the correlation analyses.

If not separated by condition, the analyses reveal a medium-sized and significant correlation between students' perceived authenticity and their situational interest, but no significant correlation between students' perceived authenticity and their performance on the knowledge test. Per condition, however, the correlation analyses reveal no significant correlation between PF students' perceived authenticity and their situational interest, but a small-sized and significant correlation between PF students' perceived authenticity and their performance on the knowledge test. In the DI condition, the results demonstrate significant and medium to large-sized correlations between students' perceived authenticity and their situational interest, but no significant correlation between their perceived authenticity and their knowledge-test performance. Thus, overall, our results of the correlation analyses partly support our H3 (only in DI condition) and partly our H4 (only in PF condition).

² Please note that we investigated the effectiveness of PF for students' knowledge-test performance in a parallel study on the same dataset (see Nachtigall et al., 2020), where we report the following descriptive statistics for students' performance on the knowledge test: PF: *M* = 9.61 (*SD* = 3.62); DI: *M* = 10.07 (*SD* = 3.87).

Fig. 2 Results of the mediation analysis with triggered situational interest as outcome variable

Fig. 3 Results of the mediation analysis with maintained situational interest as outcome variable

Table 9 Linear model of predictors of students' perceived authenticity of the learning activity

Perceived authenticity...	<i>b</i>	<i>SE B</i>	β	<i>p</i>
Of the method	.21	.08	.23	.012
Of the materials	-.07	.10	-.07	.479
Of the content	.30	.08	.33	<.001
Of the location	.01	.06	.02	.841
Of the instructor	.07	.06	.10	.270
In terms of experienced curiosity and excitement	.16	.06	.19	.009
Of the innovation	-.05	.05	-.08	.265

$R^2 = 0.33$

Mediation effects

To further examine whether the authenticity level of the learning activity has an indirect effect—mediated by students' perceived authenticity—on situational interest (H3), we conducted two mediation analyses. We used condition as a predictor variable (X), students' perceived authenticity of the learning activity as a mediator (M) variable, and their situational interest (either catch or hold) as an outcome variable (Y). As depicted in Figs. 2 and 3, the mediation analyses reveal no indirect effect (which is indicated by the bootstrap confidence intervals that cross zero) of the authenticity level of the learning activity neither on triggered nor on maintained situational interest through perceived authenticity.

Table 10 Descriptive statistics for students' perceived authenticity of different features of the learning setting

Perceived authenticity...	PF ($n=80$) M (SD)	DI ($n=72$) M (SD)
Of the method	3.44 (0.63)	3.46 (0.83)
Of the materials	3.48 (0.67)	3.48 (0.64)
Of the content	3.70 (0.65)	3.87 (0.81)
Of the location	3.49 (0.88)	3.26 (1.07)
Of the instructor	3.43 (0.94)	3.09 (0.93)
In terms of experienced curiosity and excitement	3.16 (0.74)	3.14 (0.85)
Of the innovation	1.99 (1.03)	1.94 (0.93)

Research question 2

To explore whether other features of the learning setting (than the learning activity, or in other words, the method that students used to work on a problem) affected students' perceived authenticity of the learning activity (RQ2), we conducted a multiple regression analysis. We used perceived authenticity of the learning activity as a dependent variable and perceived authenticity of different features of the learning setting as predictor variables. As this analysis is an exploratory analysis, we did not test concrete hypotheses. Using an unadjusted significance level of $\alpha=.05$, the results of the regression analysis (see Table 9) show that students' perceived authenticity of the *content*, the *method*, and their perceived authenticity in terms of *experienced curiosity and excitement* significantly predicted their perceived authenticity of the learning activity. Using a lower and more stricter significance level of $\alpha=.01$, only the content and the experienced excitement significantly predicted students' perceived authenticity of the learning activity.

The descriptive statistics with regard to students' perceived authenticity of the different features of the learning setting are depicted in Table 10. The descriptive statistics show that students' perceived authenticity was highest with regard to the content in both conditions.

It should be noted that a principal axis factor analysis with oblique rotation (direct oblimin) did not confirm the separation of the seven different scales of perceived authenticity. Instead, the factor analysis extracted four factors with an eigenvalue of 1 or higher. Factor 1 explained 26% of the variance and united the items assessing students' perceived authenticity of the method, the materials, and of the content. Factor 2, which explained 9% of the variance, united the items measuring students' perceived authenticity of the location and of the instructor. Factor 3 accounted for 6% of the variance and united all items related to students' experienced curiosity and excitement. The innovation item loaded on factor 4 and accounted for 5% of the variance. The two new scales showed the following internal consistencies (i.e., Cronbach's alphas): .82 for factor 1 (i.e., method, materials, and content) and .81 for factor 2 (i.e., location and instructor). However, as we were interested in exploring the role of different features of an authentic learning setting (other than the method) for students' perceived authenticity of the learning activity, we conducted our exploratory analyses reported here with the original seven scales that build on the theoretical authenticity model from Betz et al. (2016).

Discussion

OSLs aim at authentically communicating scientific ways of thinking and working by engaging students in authentic learning activities that emulate processes of scientific inquiry and that situate learners in the role of a scientist (e.g. Glowinski & Bayrhuber, 2011). Also, theories of authentic learning emphasize the role of certain learning activities for achieving an authentic contextualization of learning experiences (e.g. Herrington & Oliver, 2000). However, OSL research has not yet focused on investigating whether the involvement of students in activities that are assumed to authentically simulate research practices indeed has an impact on students' perceived authenticity and further motivational and cognitive learning outcomes. A learning activity that emulates scientific inquiry processes and thus is characterized as an authentic learning activity is PF (Kapur & Toh, 2015). In PF, students are asked to collaboratively explore solution ideas to a complex problem before they have to falsify³ their solutions during a delayed instruction phase. We investigated whether PF as an authentic learning activity has an impact on students' perceived authenticity and situational interest in an OSL for social sciences (RQ1). For this purpose, we compared PF with problem solving followed by instruction to DI with instruction followed by problem solving. DI can be characterized as less authentic than PF as it promotes the conceptualization of scientific inquiry being nothing else but a simple application of instructions. We moreover investigated the impact of students' perceived authenticity of different features of the learning setting on their perceived authenticity of the learning activity (RQ2).

Although the design of the present study differs from the design of our previous study, the findings with regard to RQ1 mostly replicate the results of our previous study (see Table 11). Specifically, as we did not find an effect of the authenticity level of the learning activity on students' perceived authenticity and situational interest in our previous study, we hypothesized that students probably lack knowledge about the research methods and scientific practices within the social sciences (Nachtigall et al., 2018). In the present study, we therefore introduced different research methods and a typical research process to students in both conditions. We further emphasized that experiences of failure and uncertainty are often an inherent part of scientific-inquiry processes. Nevertheless, the results of the present study demonstrated that the authenticity level of the learning activity (i.e., PF versus DI) did not have an effect on students' perceived authenticity of the learning activity (against H1) nor on their situational interest (against H2). Moreover, students' perceived authenticity of the learning activity positively correlated with their situational interest (H3), but only in the DI condition. Students' perceived authenticity also correlated with their performance on the knowledge test (H4), but only in the PF condition. Our previous study demonstrated a positive correlation between perceived authenticity and students' situational interest and no correlation between perceived authenticity and students' knowledge-test performance in both conditions (Nachtigall et al., 2018). However, the fact that the present study mostly replicated the findings of our previous study suggests that

³ As described in Nachtigall et al. (2020), during the collaborative problem-solving phase, 93% of the PF student groups interpreted correlational evidence as causal, meaning that these groups stated that a correlation study could reveal causal evidence. In fact, none of the PF groups were able to solve the problem canonically during the initial problem-solving phase. Specifically, PF students reached on average 3.18 out of 8 points and a maximum of 6 points. Hence, in the present study, all PF students had to falsify their solution ideas during delayed instruction.

Table 11 Comparison of the results between our previous and present study

Hypotheses	Previous study (N = 212) (see Nachtigall et al., 2018)	Present study (N = 152)
Effect of learning activity (PF vs. DI) on perceived authenticity	$\eta_p^2 = .003$; 90% CI [.00; .03]	$\eta_p^2 = .019$; 90% CI [.00; .07]
Effect of learning activity (PF vs. DI) on situational interest (catch)	$\eta_p^2 = .003$; 90% CI [.00; .03]	$\eta_p^2 = .000$; 90% CI [.00; .02]
Effect of learning activity (PF vs. DI) on situational interest (hold)	$\eta_p^2 = .009$; 90% CI [.00; .04]	$\eta_p^2 = .001$; 90% CI [.00; .02]
Correlation between perceived authenticity and situational interest (catch)	$r = .37, p < .001$	$r = .30, p < .001$
Correlation between perceived authenticity and situational interest (hold)	$r = .30, p < .001$	$r = .38, p < .001$
Correlation between perceived authenticity and performance on knowledge test	$r = .02, p = .75$	$r = .14, p = .08$

introducing different research methods and features of a research process within the social sciences to students did seemingly not lead to the expected beneficial effects, namely that PF students would perceive their learning activity as more authentic and interesting than DI students. The descriptive statistics even show that students' reported perceived authenticity did not differ between the present study ($M=3.22$, $SD=0.66$) and our previous study ($M=3.20$, $SD=0.61$). With respect to students' situational interest, the descriptive statistics point to small differences between the present study (catch: $M=3.15$, $SD=0.70$; hold: $M=2.87$, $SD=0.71$) and our previous study (catch: $M=2.97$, $SD=0.83$; hold: $M=2.49$, $SD=0.68$). At the same time, these similarities also suggest that our introduction of different research methods and features of research processes within the social sciences did not harm students' perceived authenticity of the following learning activities. Specifically, one could argue that this brief instruction might have primed students to perceive the following learning activities as less authentic than the introduced research activities (e.g. collecting and analyzing data) of social scientists. But the comparison of the findings between our previous and present study does not support this assumption.

A potential reason for the non-effect of the authenticity level of the learning activity on students' perceived authenticity of the learning activity (H1) could be related to the high authenticity level of other features of the learning setting that may have outweighed a potential effect of the authenticity level of the learning activity. As the results with respect to our RQ2 suggest, students in both conditions perceived the content of the OSL project—which was the same in both conditions—as highly authentic. In addition, the perceived authenticity of the content significantly predicted students' perceived authenticity of the learning activity. The perceived authenticity of the method also predicted students' perceived authenticity of the learning activity (but only when using an unadjusted significance level of $\alpha=.05$). This finding is not surprising, as the method that students used to work on a problem is closely related to the learning activity. A further significant predictor of students' perceived authenticity of the learning activity was their experienced excitement and curiosity. Surprisingly, students in both conditions reported similarly high experienced excitement and curiosity. One could have expected that PF students would experience higher excitement and curiosity due to the delayed instruction phase and, consequently, due to students' potential uncertainty about the quality of their solution ideas during the initial problem-solving phase. However, as both PF and DI students reported high excitement and curiosity, the entire learning situation, which was novel to students in both conditions and that differed from their usual learning experiences in school (i.e., new instructor, new location, new learning topic, etc.), might have had an influence on students' experienced excitement and curiosity. Hence, we hypothesize that the different features of an authentic and/or novel learning setting are interrelated, such that the authenticity level of one feature (e.g. the content) may affect students' perceived authenticity of another feature (e.g. the learning activity). This hypothesized interrelatedness of different features of an authentic learning setting is further supported by the findings of our factor analysis of the items assessing students' perceived authenticity of seven different features of the learning setting. As the results of this factor analysis demonstrated, the two features instructor and location on the one hand and the three features method, material, and content on the other hand could not be separated from each other.

A similar reason could be associated with the non-effect of the authenticity level of the learning activity on students' situational interest (H2). Specifically, further or even other factors than the authenticity level of the learning activity might have affected students' situational interest. This assumption is also supported by our correlation analyses showing no association between PF students' perceived authenticity of the learning activity and

their situational interest (H3). Hidi and Renninger (2006), for instance, list several features of the learning setting that are expected or that have been demonstrated to foster students' development of situational interest. Some of these features are collaborative group work, computer work, or instructional materials with surprising information (Hidi & Renninger, 2006). In the present study, students in both conditions were asked to collaboratively solve a problem and to use a laptop to write down their solution ideas. Moreover, it is likely that some of the information given during instruction (e.g. that scientists also do mistakes or that research may reveal inconsistent findings) were surprising to students in both conditions. Hence, we hypothesize that the often anticipated effects of an authentically designed learning setting on different motivational and cognitive learning outcomes (e.g. Betz et al., 2016; Brown et al., 1989; Lepper, 1988; Newman & Wehlage, 1993), such as students' situational interest, may be overrated. Consequently, there is a need to further investigate the effectiveness of authentic learning settings and to assess students' perceived authenticity. It, otherwise, remains unclear whether students indeed perceive an authentically designed learning setting as authentic and whether students' perceived authenticity is in fact associated with, for instance, their situational interest.

In summary, our findings, which mostly replicate the results of our previous study (see Table 11), suggest that the intended authenticity level as well as students' perceived authenticity of a learning activity seem to play a less important role for students' situational interest and knowledge acquisition than hypothesized by characterizations of OSLs and theories of authentic learning more generally. In light of our findings, it seems necessary to further examine theoretical assumptions related to the effectiveness of authentically contextualized learning. At the same time, our findings also suggest that investigating the role of certain features hypothesized to determine the authenticity of learning settings, such as the learning activity, appears difficult due to the potentially interrelatedness of the features of an authentic learning setting. Future research should focus on investigating this interrelatedness as well as the effectiveness of authentic learning settings. For this purpose, it seems necessary and promising to assess students' perceived authenticity of different features of the learning setting and to examine the impact of students' perceived authenticity on different learning outcomes. The present study was an initial attempt to pinpoint this direction.

Limitations

Two limitations of the present study relate to the interpretation of the findings: Firstly, as the study took place in an OSL that is usually visited by whole classes, we had to implement a quasi-experimental design. Thus, we were unable to randomly assign students individually to the conditions, whereby our two test groups may be not equivalent. Secondly, the analyses revealed non-effects of the authenticity level of the learning activity on students' perceived authenticity and situational interest. Non-effects revealed by null hypothesis significance tests (NHST), such as the F-test, are difficult to interpret, as a NHST examines whether a null hypothesis can be rejected (which was in line with our expectations) and not whether a null hypothesis can be accepted. To address these two limitations, we conducted further analyses. Namely, multilevel regression analyses which account for the hierarchical structure of our data (i.e., individual students clustered in whole classes) and Bayesian variance analyses which allow to test

null hypotheses. The results of these additional analyses (see Online Resources 1 and 2) are in line with the results of the variance analyses reported in this paper.

Two further limitations of the present study, which should be addressed in future research, relate to the missing assessment of students' (1) knowledge about social science research methods in a pretest and (2) perceived authenticity of the learning activity after the first learning phase. We did not administer a pretest assessing students' knowledge about social science research methods, because such a pretest could have activated relevant prior knowledge in students. Consequently, the beneficial effects hypothesized to underlie the initial problem-solving phase of PF could have been negated (see Newman & DeCaro, 2018; see also Nachtigall et al., 2020). However, the design of the present study partly built on the assumption that secondary school students lack knowledge about social science research methods and, thus, are unable to accurately judge the authenticity of learning activities that try to emulate processes of scientific inquiry within the social sciences. To further investigate this assumption, future research should try to find ways for assessing and taking students' prior knowledge about processes of scientific inquiry into account. Also, it could be interesting to investigate the development of students' perceived authenticity between different learning phases. Regarding the present study, one could assume that PF students might have perceived the initial problem-solving phase as more authentic than the delayed (and possibly school-like) instruction phase, or that the instruction phase even harmed students' overall perceived authenticity of the learning activity. On the contrary, PF students might have perceived the problem-solving phase as an ordinary group-work activity that they know from school and, thus, as less authentic than the instruction given by a prospective scientist. As students in both conditions reported a rather high perceived authenticity of the method, and as DI students even reported slightly higher perceived authenticity than PF students, it is likely that the instruction phase did not harm their perceived authenticity. Nevertheless, further research is required in order to investigate whether students' perceived authenticity differs between different learning phases.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1007/s11251-021-09556-3>.

Acknowledgements The content of this manuscript has been partly published as part of the dissertation thesis of the first author (Nachtigall, 2019). This research was funded by the Graduate School of Educational Studies of Ruhr-University Bochum (RUB) within the doctoral program on science education in out-of-school labs. We thank the members for their feedback on the design of this study. We would especially like to thank Dr. Katja Serova from the Institute of Educational Research at RUB who provided insight and expertise that greatly supported the research presented in this paper. We are very thankful to the team of the Alfried Krupp-Schülerlabor (out-of-school lab) for their cooperation, and to the participating schools and teachers for their organizational efforts. We thank our student research assistants for their help in collecting the data. Moreover, we thank our colleagues at the Educational Psychology Research Group at RUB for their help while writing this paper. We also thank the anonymous reviewers for critically reading our manuscript and for providing suggestions that helped to substantially improve and clarify this paper.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval Data collection followed the standards of the German Psychological Society (DGPs). Given the type of data collected (i.e., anonymous data), no ethical approval was considered necessary according to the guidelines at that time.

Informed consent Written informed parental consent was given for all participants.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aberson, C. (2002). Interpreting null results: Improving presentation and conclusions with confidence intervals. *Journal of Articles in Support of the Null Hypotheses*, 1(3), 36–42.
- Beier, M. E., Kim, M. H., Saterbak, A., Leautaud, V., Bishnoi, S., & Gilberto, J. M. (2019). The effect of authentic project-based learning on attitudes and career aspirations in STEM. *Journal of Research in Science Teaching*, 56(1), 3–23.
- Betz, A. (2018). Der Einfluss der Lernumgebung auf die (wahrgenommene) Authentizität der linguistischen Wissenschaftsvermittlung und das Situationale Interesse von Lernenden (The influence of the learning environment on learners' (perceived) authenticity of science communication and on their situational interest). *Unterrichtswissenschaft*, 46(3), 261–278.
- Betz, A., Flake, S., Mierwald, M., & Vanderbeke, M. (2016). Modelling authenticity in teaching and learning contexts. A contribution to theory development and empirical investigation of the construct. In C.-K. Looi, J. Polman, U. Cress, & P. Reimann (Eds.), *Transforming learning, empowering learners: The international conference of the learning sciences (ICLS) 2016* (Vol. 2, pp. 15–818). International Society of the Learning Sciences.
- Braund, M., & Reiss, M. (2006). Towards a more authentic science curriculum: The contribution of out-of-school learning. *International Journal of Science Education*, 28(12), 1373–1388.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32–42.
- Chalmers, A. F. (2013). *What is this thing called science?* Hackett Publishing.
- Chang, C. W., Lee, J. H., Wang, C. Y., & Chen, G. D. (2010). Improving the authentic learning experience by integrating robots into the mixed-reality environment. *Computers & Education*, 55(4), 1572–1578.
- Chen, S. Y., Feng, Z., & Yi, X. (2017). A general introduction to adjustment for multiple comparisons. *Journal of Thoracic Disease*, 9(6), 1725–1729.
- Cho, Y. H., Caleon, I. S., & Kapur, M. (2015). *Authentic problem solving and learning in the 21st century*. Springer.
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284–290.
- Damerau, K. (2012). *Molekulare und Zell-Biologie im Schülerlabor—Fachliche Optimierung und Evaluation der Wirksamkeit im BeLL Bio (Bergisches Lehr-Lern-Labor Biologie). (Molecular and cell biology in an out-of-school lab—Professional optimization and evaluation of the effectiveness of the out-of-school lab at university of Wuppertal)*. [Dissertation]. [Wuppertal, Germany]: University of Wuppertal.
- Engeln, K. (2004). *Schülerlabors: Authentische aktivierende Lernumgebungen als Möglichkeit, Interesse an Naturwissenschaften und Technik zu wecken (Out-of-school labs: Authentic and activating learning environments as possibility to arouse interest in natural sciences and technology)*. Logos.
- Euler, M. (2004). Quality development: Challenges to physics education. In M. Michelini (Ed.), *Quality development in teacher education and training* (pp. 17–30). Forum.
- Firestein, S. (2015). *Failure: Why science is so successful*. Oxford University Press.
- Fougt, S. S., Misfeldt, M., & Shaffer, D. W. (2019). Realistic authenticity. *Journal of Interactive Learning Research*, 30(4), 477–504.
- Garner, N., & Eilks, I. (2015). The expectations of teachers and students who visit a non-formal student chemistry laboratory. *Eurasia Journal of Mathematics, Science & Technology Education*, 11(5), 1197–1210.

- Glowinski, I. (2007). *Schülerlabore im Bereich Molekularbiologie als Interesse fördernde Lernumgebungen (Out-of-school laboratories in the area of molecular biology as interest-supporting learning environments)*. [Dissertation]. [Kiel, Germany]: University of Kiel.
- Glowinski, I., & Bayrhuber, H. (2011). Student labs on a university campus as a type of out-of school learning environment: Assessing the potential to promote students' interest in science. *International Journal of Environmental and Science Education*, 6(4), 371–392.
- Gulikers, J. T. M., Bastiaens, T. J., & Kirschner, P. A. (2006). Authentic assessment, student and teacher perceptions: The practical value of the five dimensional-framework. *Journal of Vocational Education and Training*, 58(3), 337–357.
- Gulikers, J. T. M., Bastiaens, T. J., & Martens, R. L. (2005). The surplus value of an authentic learning environment. *Computers in Human Behavior*, 21(3), 509–521.
- Gulikers, J. T., Kester, L., Kirschner, P. A., & Bastiaens, T. J. (2008). The effect of practical experience on perceptions of assessment authenticity, study approach, and learning outcomes. *Learning and Instruction*, 18(2), 172–186.
- Herrington, J., & Oliver, R. (2000). An instructional design framework for authentic learning environments. *Educational Technology Research and Development*, 48(3), 23–48.
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, 41(2), 111–127.
- Hod, Y., & Sagy, O. (2019). Conceptualizing the designs of authentic computer-supported collaborative learning environments in schools. *International Journal of Computer-Supported Collaborative Learning*, 14(2), 143–164.
- Hodson, D. (1999). Going beyond cultural pluralism: Science education for sociopolitical action. *Science Education*, 83(6), 775–796.
- Itzek-Greulich, H., Flunger, B., Vollmer, C., Nagengast, B., Rehm, M., & Trautwein, U. (2015). Effects of a science center outreach lab on school students' achievement—Are student lab visits needed when they teach what students can learn at school? *Learning and Instruction*, 38, 43–52.
- Itzek-Greulich, H., Flunger, B., Vollmer, C., Nagengast, B., Rehm, M., & Trautwein, U. (2017). Effectiveness of lab-work learning environments in and out of school: A cluster randomized study. *Contemporary Educational Psychology*, 48, 98–115.
- Itzek-Greulich, H., & Vollmer, C. (2017). Emotional and motivational outcomes of lab work in the secondary intermediate track: The contribution of a science center outreach lab. *Journal of Research in Science Teaching*, 54(1), 3–28.
- Kapur, M. (2015). Learning from productive failure. *Learning: Research and Practice*, 1(1), 51–65.
- Kapur, M., & Bielaczyc, K. (2012). Designing for productive failure. *The Journal of the Learning Sciences*, 21(1), 45–83.
- Kapur, M., & Toh, L. (2015). Learning from productive failure. In Y. H. Cho, I. S. Caleon, & M. Kapur (Eds.), *Authentic problem solving and learning in the 21st century* (pp. 213–227). Springer.
- Knogler, M., Harackiewicz, J. M., Gegenfurter, A., & Lewalter, D. (2015). How situational is situational interest? Investigating the longitudinal structure of situational interest. *Contemporary Educational Psychology*, 43, 39–50.
- Krapp, A. (2002). Structural and dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective. *Learning and Instruction*, 12(4), 383–409.
- Lepper, M. R. (1988). Motivational considerations in the study of instruction. *Cognition and Instruction*, 5(4), 289–309.
- Lewalter, D., & Geyer, C. (2009). Motivationale Aspekte von schulischen Besuchen in naturwissenschaftlich-technischen Museen (Motivational aspects of school visits to natural science and technical museums). *Zeitschrift Für Erziehungswissenschaft*, 12(1), 28–44.
- Lewalter, D., & Willems, A. S. (2009). Die bedeutung des motivationsrelevanten erlebens und des individuellen fachinteresses für das situationale interesse im mathematikunterricht (The relevance of emotional experiences and subject-related interest for situational interest in mathematics). *Psychologie in Erziehung Und Unterricht*, 56, 243–257.
- Loibl, K., Roll, I., & Rummel, N. (2017). Towards a theory of when and how problem solving followed by instruction supports learning. *Educational Psychology Review*, 29(4), 693–715.
- Luehmann, A. L. (2009). Accessing resources for identity development by urban students and teachers: Foregrounding context. *Cultural Studies of Science Education*, 4(1), 51–66.
- Mierwald, M., Lehmann, T., & Brauch, N. (2018). Zur Veränderung epistemologischer Überzeugungen im Schülerlabor: Authentizität von Lernmaterial als Chance der Entwicklung einer wissenschaftlich angemessenen Überzeugungshaltung im Fach Geschichte? (Changing epistemological beliefs in student labs: Authentic learning materials as a chance to foster the development of academically adequate beliefs in the domain of history?). *Unterrichtswissenschaft*, 46(3), 279–297.

- Nachtigall, V. (2019). *Failing just as scientists do? The effectiveness of Productive Failure for learning in an out-of-school lab for social sciences*. (Doctoral dissertation, Ruhr-Universität Bochum, Germany). Institutional Repository at the Ruhr-Universität Bochum. <https://doi.org/10.13154/294-6617>
- Nachtigall, V., Rummel, N., & Serova, K. (2018). Authentisch ist nicht gleich authentisch—Wie Schülerinnen und Schüler die Authentizität von Lernaktivitäten im Schülerlabor einschätzen (Authentic does not equal authentic—How students evaluate the authenticity of learning activities in an out-of-school lab). *Unterrichtswissenschaft*, 46(3), 299–319.
- Nachtigall, V., Serova, K., & Rummel, N. (2020). When failure fails to be productive: Probing the effectiveness of productive failure for learning beyond STEM domains. *Instructional Science*, 48(6), 651–697.
- Newman, P., & DeCaro, M. (2018). How much support is optimal during exploratory learning? In *Proceedings of the 40th annual conference of the cognitive science society*. Cognitive Science Society.
- Newmann, F. M., & Wehlage, G. G. (1993). Five standards of authentic instruction. *Educational Leadership*, 50(7), 8–12.
- Pauly, Y. (2012). Was sind und zu welchem Zweck brauchen wir geisteswissenschaftliche Schülerlabore? (What are and for what purpose do we need out-of-school labs on humanities and social sciences)]. *Handbuch Wissenschaftskommunikation* (pp. 205–210). VS Verlag für Sozialwissenschaften.
- Pawek, C. (2009). *Schülerlabore als interessesfördernde außerschulische Lernumgebungen für Schülerinnen und Schüler aus der Mittel- und Oberstufe (Out-of-school labs as interest-promoting out-of-school learning environments for students of secondary schools)*. [Dissertation]. [Kiel, Germany]: University of Kiel.
- Polman, J. L. (2012). Trajectories of participation and identification in learning communities involving disciplinary practices. In D. Yun Dai (Ed.), *Design research on learning and thinking in educational settings: Enhancing intellectual growth and functioning* (pp. 225–242). Routledge.
- Rost, D. H., Sparfeldt, J. R., & Buch, S. R. (2008). Kann denn Kürze Sünde sein? Erfassung schulfachspezifischer Interessen mit nur einem Item (Can brevity be a sin? Assessment of subject-related interests by using only one item). In F. Hofmann, C. Schreiner, & J. Thonhauser (Eds.), *Qualitative und quantitative Aspekte. Zu ihrer Komplementarität in der erziehungswissenschaftlichen Forschung* (pp. 225–238). Waxmann.
- Rule, A. C. (2006). Editorial: The components of authentic learning. *Journal of Authentic Learning*, 3(1), 1–10.
- Sauter, M., Uttal, D. H., Rapp, D. N., Downing, M., & Jona, K. (2013). Getting real: The authenticity of remote labs and simulations for science learning. *Distance Education*, 34(1), 37–47.
- Scharfenberg, F. J., & Bogner, F. X. (2014). Outreach science education: Evidence-based studies in a Gene Technology Lab. *Eurasia Journal of Mathematics, Science & Technology Education*, 10(4), 329–341.
- Scharfenberg, F. J., Bogner, F. X., & Klautke, S. (2007). Learning in a gene technology laboratory with educational focus: Results of a teaching unit with authentic experiments. *Biochemistry and Molecular Biology Education*, 35(1), 28–39.
- Schmidt, S. (2016). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, 13(2), 90–100.
- Schüttler, T., Watzka, B., Girwidz, R., & Ertl, B. (2021). Die Wirkung der Authentizität von Lernort und Laborgeräten auf das situationale Interesse und die Relevanzwahrnehmung beim Besuch eines naturwissenschaftlichen Schülerlabors (Effects of an authentic location and laboratory equipment for the situational interest and the perception of content relevance when visiting an out-of-school science lab). *Zeitschrift Für Didaktik Der Naturwissenschaften*. <https://doi.org/10.1007/s40573-021-00128-z>
- Shaffer, D. W., & Resnick, M. (1999). “Thick” authenticity: New media and authentic learning. *Journal of Interactive Learning Research*, 10(2), 195–216.
- Smeds, P., Jeronen, E., & Kurppa, S. (2015). Farm education and the value of learning in an authentic learning environment. *International Journal of Environmental and Science Education*, 10(3), 381–404.
- Sparfeldt, J. R., Rost, D. H., & Schilling, S. R. (2004). Schulfachspezifische Interessen—Ökonomisch gemessen. (An economical assessment of subject-specific interests in grammar school students). *Psychologie in Erziehung Und Unterricht*, 51, 213–250.
- Stamer, I., David, M. A., Höffler, T., Schwarzer, S., & Parchmann, I. (2021). Authentic insights into science: Scientific videos used in out-of-school learning environments. *International Journal of Science Education*. <https://doi.org/10.1080/09500693.2021.1891321>
- Waldow, F. (2009). What PISA did and did not do: Germany after the ‘PISA-shock.’ *European Educational Research Journal*, 8(3), 476–483.
- Wilde, M., Hußmann, J. S., Lorenzen, S., Meyer, A., & Randler, C. (2012). Lessons with living harvest mice: An empirical study of their effects on intrinsic motivation and knowledge acquisition. *International Journal of Science Education*, 34(18), 2797–2810.

Wirth, J., Raimann, J., & Fleischer, J. (2017). *Wahrnehmung von Authentizität der Wissenschaftsvermittlung. (Perceived authenticity of science communication)*. Poster presented at the conference “scientific investigation and evaluation of out-of-school labs”.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.