International Journal of Science Education
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tsed20

Promoting Experimental Problem-solving Ability in Sixth-grade Students Through Problem-oriented Teaching of Ecology: Findings of an intervention study in a complex domain

Frank Roesch[a], Josef Nerb[b] & Werner Riess[c]
[a] Biology Department, Institute of Natural Sciences and Technology, University of Education Ludwigsburg, Ludwigsburg, Germany
[b] Institute of Psychology, University of Education Freiburg, Freiburg, Germany
[c] Institute of Biology and its Didactics, University of Education Freiburg, Freiburg, Germany
Published online: 14 Jan 2015.

International Journal of Science Education, 2015
http://dx.doi.org/10.1080/09500693.2014.1000427

Promoting Experimental Problem-solving Ability in Sixth-grade Students Through Problem-oriented Teaching of Ecology: Findings of an intervention study in a complex domain

Frank Roesch[a*], Josef Nerb[b] and Werner Riess[c]
[a]*Biology Department, Institute of Natural Sciences and Technology, University of Education Ludwigsburg, Ludwigsburg, Germany;* [b]*Institute of Psychology, University of Education Freiburg, Freiburg, Germany;* [c]*Institute of Biology and its Didactics, University of Education Freiburg, Freiburg, Germany*

Our study investigated whether problem-oriented designed ecology lessons with phases of direct instruction and of open experimentation foster the development of cross-domain and domain-specific components of *experimental problem-solving ability* better than conventional lessons in science. We used a paper-and-pencil test to assess students' abilities in a quasi-experimental intervention study utilizing a pretest/posttest control-group design ($N = 340$; average performing sixth-grade students). The treatment group received lessons on *forest ecosystems* consistent with the principle of education for sustainable development. This learning environment was expected to help students enhance their ecological knowledge and their theoretical and methodological experimental competencies. Two control groups received either the teachers' usual lessons on forest ecosystems or non-specific lessons on other science topics. We found that the treatment promoted specific components of experimental problem-solving ability (generating epistemic questions, planning two-factorial experiments, and identifying correct experimental controls). However, the observed effects were small, and awareness for aspects of higher ecological experimental validity was not promoted by the treatment.

Keywords: *Biology Education; Inquiry-based Teaching; Experimentation; Problem-solving; Ecology Education; Context-based Learning*

*Corresponding author. Biology Department, Institute of Natural Sciences and Technology, University of Education Ludwigsburg, Reuteallee 46, Ludwigsburg D-71634, Germany. Email: roesch@ph-ludwigsburg.de

Introduction

Experimentation is considered to be one of the most important methods for acquiring insight into the area of scientific thinking and inquiry activities (Carey, Evans, Honda, Jay, & Unger, 1989). Thus, the promotion of competencies that facilitate the independent planning, implementation, analysis, and evaluation of experiments is an important educational goal of science education. School achievement studies, such as the Third International Mathematics and Science Study in the 1990s and the Program for International Student Assessment, have long borne testimony to the relatively low level of problem-solving ability of German students with respect to the acquisition of sophisticated natural science knowledge (Prenzel, Sälzer, Klieme, & Köller, 2013). The consequence of these studies was that, in 2004, the National Conference of German Ministers of Education in Germany introduced an area of competence for teachers to target in *epistemic methods/scientific inquiry*, with several inquiry skills related to experimentation in current curricula (KMK, 2005). As a result, German science teachers are required to promote both learning science content and learning to do science.

A focus of German life science curricula is environmental awareness and the associated systems thinking skills required for making informed and wise decisions on issues that affect sustainable economic and environmental development (Riess & Mischo, 2010). Such competencies are often referred to as 'ecological literacy' or 'environmental literacy'. While some authors see environmental literacy just as a predecessor to ecological literacy (e.g. Cutter-Mackenzie & Smith, 2003), others do not identify any differences between these two concepts (e.g. Orr, 1992), or they consider ecological knowledge as one component of environmental literacy (e.g. Morrone, Mancl, & Carr, 2001). We propose the use of the term 'ecological literacy' when ecological aspects of sustainability are at the focus, and 'environmental literacy', when social, economic, and ecological issues are equally concerned. In our study, we have laid the focus on ecological literacy; above all, we wanted the participating students to investigate ecological relationships by learning about and using scientific reasoning and inquiry procedures. Thus, we did not implement interdisciplinary lessons: social and economic issues only played a subordinate part in our treatment context. Ecological literacy might be acquired through systematic observation, measurement, and experimentation (e.g. Berkowitz, 1997; McBride, Brewer, Berkowitz, & Borrie, 2013). Thus, an ecologically literate person understands environmental realities by specifically identifying their cause-and-effect relationships between elements of the ecosystems (McBride et al., 2013).

Because German science teachers are allowed to design their own lesson plans and to arrange teaching–learning environments by combining competencies according to the national output standards with any domain, respectively, learning context, some central questions must be answered: How and when can experimental inquiry skills be promoted effectively? (Hofstein & Lunetta, 2004; Lunetta, Hofstein, & Clough, 2007). Would it be possible to enhance the experimental problem-solving ability in a complex domain such as *ecology* which plays an important part in our everyday lives? Is problem-oriented learning in such an authentic context beneficial to foster problem-solving skills? To date, it has not yet been clarified whether experimentation in the science classroom should be promoted more in isolation, removed from specific challenging scientific content, or whether the contextual embedding of competence development into more complex content can generate particular opportunities for learning processes (Keselman, 2003). Thus, whether and how experimental problem-solving ability can be promoted in complex domains that are part of the natural science curricula, such as ecology, is an important research issue and is highly relevant for science teaching.

To advance the discussion on these issues, we reviewed the research literature on promoting experimental problem-solving ability in the classroom. Our focus was on the extent that complexity varies both as a function of the domain (=topic of instruction) under scrutiny (Keselman, 2003) and as a function of the required cognitive demands within a specific learning context (Kirschner, Sweller, & Clark, 2006) (e.g. particularly challenging learning activities, respectively, the topic in isolation or within the frame of sustainable development). We designed our study to scrutinize the effectiveness of a specially designed teaching module on this topic and to deduce practical implications and research desiderata.

Theoretical Background and State of Research

Experimental Problem-solving Ability

Experimenting—a demanding inquiry 'tool' to solve scientific problems. Experimentation is seen as a complex problem-solving process (Hammann, Phan, Ehmer, & Grimm, 2008), comprising various operations requiring cognitive, methodical, and technical skills (Carey et al., 1989; Germann, Aram, Odom, & Burke, 1996; Mayer, Grube, & Möller, 2008)—as well as the metacognitive ability to self-regulate (Thillmann, Künsting, Wirth & Leutner, 2009).

Competence models for experimentation skills. Many approaches in science didactics for fostering experimentation skills are based on the *Scientific Discovery as Dual Search* (SDDS) model proposed by Klahr (2000). The SDDS model describes the phases and cognitive activities in scientific inquiry in detail, helps to analyze students' levels in experimentation competencies, and takes into account the effects of inadequate conceptions and inappropriate strategies. It has been shown to be useful as a basis for promoting and assessing science inquiry skills (Ehmer, 2008; Ganser & Hammann, 2009; Hof, 2011; Klahr & Nigam, 2004; Neber & Anton, 2008).

Being based on a psychological theory of human problem-solving (Newell & Simon, 1972), the SDDS model considers experimentation as being a problem-solving process within two spaces: a hypothesis space, within which domain-specific prior knowledge can be used to generate theory-based hypotheses that could explain phenomena, and an experiment space that consists of all possible ways to conduct experiments within the domain. Thus, selecting a hypothesis and choosing an experimental design to test the hypothesis are seen as the main tasks of the problem solver.

Klahr (2000) emphasized three major problem-solving phases: in the first, the problem solver searches the hypothesis space; afterwards, he searches the experiment space, and finally evaluates the evidence. These processes are interconnected and consist of different subprocesses and cognitive activities. In the second phase, adequate designs, procedures, and analytical methods are developed. The experimental design must be evaluated with respect to its internal validity. That is, before starting with actual measuring activities, effects resulting from the manipulation of the experimental system must be anticipated. Thus, contemplating imaginary experiments is an important step prior to testing a hypothesis. In the third phase of experimentation, conclusions are derived from the empirical evidence in order to accept, reject, or revise the hypothesis.

To date, different competence models based on Klahr's (2000) SDDS model have shown that many students have an unscientific understanding of the nature of science and have problems with the planning, implementation, and analysis of their experiments (Carey et al., 1989; Ehmer, 2008; Germann et al., 1996; Hammann et al., 2008; Keselman, 2003; Mayer et al., 2008). Mayer et al. (2008) and Schneider, Bullock, and Sodian (1998) found that correct understanding and use of strategies and scientific reasoning skills do improve with time.

Promotion of experimental problem-solving ability. Inquiry-based science education seems to be suitable for promoting experimental problem-solving ability (Ehmer, 2008; Keselman, 2003). Following this approach, teaching and learning processes focus on science investigation processes. Thereby, students acquire science content knowledge, and they become trained in inquiry skills and methods. Inquiry-based science education includes, among other principles, autonomous learning (Hof, 2011, pp. 31ff.): the students should not only act according to the teachers' detailed instructions, but also make decisions on their own as often as is reasonable. This aspect aims to enable the students to apply their inquiry skills and problem-solving ability independently. Moreover, according to the self-determination theory (Deci, & Ryan, 2000), more autonomy in learning processes can increase the students' intrinsic motivation—as far as they are not overly challenged by their tasks. On the other hand, autonomy in open-ended experimentation causes a more extensive need for self-regulation, compared to learning activities in guided-inquiry learning environments. A lack of guidance seems to create difficulties for our working memory, which impairs the self-regulation competencies (Kirschner et al., 2006). Hence, it is important to implement learning environments which consider both—instruction and autonomy—in balanced proportions.

Comparable intervention studies that cover the promotion of experimental competencies show different findings: for example, Ehmer (2008) observed remarkable effects in promoting procedural abilities and knowledge of scientific methods after a treatment over only four lessons. In contrast, Ganser and Hammann (2009) could not detect any effects in performance, although their didactic procedure was very similar, and their treatment had a longer duration. Both studies followed the *cognitive apprenticeship* concept, resulted in cumulative learning, and enhanced the openness of the learning environment and the autonomy of the students. However, contextualized learning did play a more important part in Ganser and Hammann's study than in Ehmer's study. Surprisingly, Hof (2011) found that a higher level of openness in experimentation can favor the enhancement of experimental problem-solving ability, compared to a treatment with more guided inquiry.

What could be the cause for the non-appearance of expected promotion effects in some intervention studies? Being engaged independently in solving experimental problems with high complexity causes a high cognitive load on working memory (Kirschner et al., 2006; see above). It seems that this condition depends both on the type of activity (e.g. computer-based simulations versus real experimentation) and on the level of openness of the experimentation activities. The high demands (a) of the learning context and (b) of the complexity of the domain are also likely to play a part (Keselman, 2003). Below, we will deal with both aspects more in detail.

Several studies show the importance of instructional support during experimentation for successful independent problem-solving; explicit instruction and guided training of scientific reasoning appear to promote cognitive and metacognitive components of experimental problem-solving ability more effectively than strictly constructivist guided discovery lessons (Kirschner et al., 2006; Klahr & Nigam, 2004). Various support methods have been tested such as (a) instructional support during teacher-led conversation in class (Socratic method) to structure, self-regulate, reflect, and discuss the inquiry process (Carey et al., 1989), (b) worked examples to get to know the steps of experimental problem-solving processes and sample problem solutions (Kirschner et al., 2006), (c) process worksheets that could help to structure the experimental inquiry process (Kirschner et al., 2006), (d) prompts in computer-based learning environments with dynamic model systems (Thillmann et al., 2009), and (e) given question stems as structured tutorial aid to generate good research questions (Neber & Anton, 2008). As a rule, these measures of instructional support only lead to success after an explicit introduction and training activities.

Are specific science content domains better suited toward fostering experimental problem-solving ability? Effective intervention studies have focused on areas of mechanics (e.g. Klahr & Nigam, 2004; Thillmann et al., 2009), acids and bases, food biochemistry (e.g. Carey et al., 1989; Ganser & Hammann, 2009), and human physiology (e.g. influences on pulse rate; Ehmer, 2008). All of these science domains are characterized by systems that are less complex than the ecological one we use. That is, the number of potential causes in stimulating an effect is limited and therefore manageable, the independent variables can be mostly well controlled, and only a few confounding variables exist or are not explicitly considered. The advantage of such low-complexity domains is that fundamental experimentation strategies and an adequate conception of the function of experiments can be relatively easily acquired. These advantages notwithstanding, some important aspects of experimentation are not involved, and thus cannot be promoted. For these reasons, we chose the domain of ecology as a test for a learning context in which various components of experimental problem-solving ability might be promoted. As far as we are aware, there have been no previous studies that use an ecological context for this purpose.

Experimentation in the Complex Domain of Ecology

Experimentation in the domain of ecology is very challenging for various reasons (Smith & Smith, 2009): first, very complex synergistic interactions exist between abiotic and biotic system elements on different hierarchical levels of living systems. This causes amazing emergent scenarios, that is, incalculable and consequently unforeseen and also less reproducible phenomena with multidimensional cause–effect relations. Second, there are a large number of potential confounding variables in ecosystems. These uncontrolled factors reduce the internal validity of experiments. Third, biological systems—particularly ecological ones—are characterized by a dynamic nature and by the complex self-regulating processes of autopoietic living systems. For these and other reasons, it is important to enhance the validity and reliability of experiments, for example, by using large sample sizes, repeated measurements, and replication. In addition, long-term effects need to be taken into account. These aspects seem to be quite complicated; even older students have difficulties with accounting for confounding variables, using appropriate durations of observations, replicating findings or repeating measurements, and with planning appropriate sample sizes (Arnold, Kremer, & Mayer, 2013). Thus, promoting experimental problem-solving ability in an ecological domain requires not only focusing on cross-domain aspects of internal validity but also on domain-specific aspects of internal, external, and ecological validity. It is an assumption of this study that evaluation of internal, external, and ecological validity could be optimally handled in an ecological domain.

To date, there have been no studies on whether awareness of these important criteria for the validity of ecological experiments can be promoted in lower school grades. This is not a trivial question; surprisingly, studies have shown that even elementary school students can be trained to understand, for example, the principles of variable control and hypothesis testing (Klahr & Nigam, 2004); and adequate understanding of multivariable causality can be enhanced already in sixth-grade students (Keselman, 2003). In view of these facts, we assumed that also the promotion of the awareness of validity aspects could be successful.

Assessing experimental problem-solving competence within an ecological context therefore requires not only measuring strategies for controlling variables, generating hypotheses about the causal relationship of variables, and decisions about the appropriateness of study designs (which are all questions of internal validity) but also has to take into account reliability issues, such as the use of appropriate sample sizes and aspects of long-term effects. To date, only few instruments have been made available to assess these competencies (e.g. Arnold et al., 2013).

Promoting Experimental Competencies in Meaningful, Authentic Contexts

As mentioned earlier, it is not yet clear what role learning context plays in the promotion of experimental problem-solving skills. In other intervention studies that used moderately demanding contexts, the focus was only on the inquiry process and skills (while the science content played a subordinate role), or the competence areas of *epistemic methods/scientific inquiry* and *scientific content knowledge* were treated equally. This is remarkable considering that didactic approaches using context-based *science-technology-society* (STS) and *socioscientific issues* (SSI) concepts are becoming more widely used in science teaching (Sadler, 2004), for example, national science education programs such as *Biology in Context* (Ganser & Hammann, 2009) and *Chemistry in Context* (Parchmann et al., 2006). These teaching programs include problem-based situated learning and also incorporate two other competence areas of *decision-making* and *communication about science* (Sadler, 2004), which were also specified as educational goals by the National Conference of German Ministers of Education (KMK, 2005).

Many studies have shown the positive effects of context-based and applications-led science education (Sadler, 2004). These studies found improvement of students' general attitudes to natural sciences and to school science, learning-motivation, argumentation skills, understanding of nature of science, and facilitation of transfer and application of competencies. The learning of science concepts and principles in context-based treatments shows effects that are at least comparable to those of conventional instruction, and sometimes even better results (Parchmann et al., 2006). These arguments speak for contextualization of teaching content in order to promote experimental problem-solving ability.

Nevertheless, contextualized science education includes a number of diverse learning goals, and thus imposes challenging demands on both instruction and learning. Some researchers have emphasized the risks of contexts with high complexity (Sadler, 2004). Such learning environments can be discouraging for students and may reduce learning effects because of high extraneous cognitive load (Kirschner et al., 2006). But problem-solving processes in experimentation require a 'didactically' sensible reduction of the complexity of phenomena and the learning environment, and a focus and specialization on central points and competencies (Muckenfuß, 1995; Parchmann et al., 2006).

Research Questions and Hypotheses

Given that open-ended and independent experimentation of students is seen as a very complex activity, the question arises whether the enhancement of its competencies should be combined with the domain of ecology in a demanding authentic context as suggested earlier. In view of the presented theoretical and empirical background, different research desiderata arose that we wanted to investigate in our study. Hence, we focused on the following issues.

(1) Does a problem-oriented teaching concept in the complex domain of ecology in a meaningful, sophisticated context promote the following components of

experimental problem-solving ability: formulation of epistemic questions, planning of nonconfounded two-factor experiments, correct handling of independent variables, and awareness of the effect of sample size and observation duration?
(2) How does our treatment affect the pupils' feeling of autonomy?

For our study, we expected a moderate effect of the teaching lessons in fostering the specific experimental problem-solving abilities; our treatment included well-demonstrated training methods (Ehmer, 2008; Hof, 2011; Neber & Anton, 2008). We assumed that sixth graders are not too young to be supported in learning about the meaning of sample size and duration of observation; experimental investigations on seed germination (Hammann et al., 2008) and the taxis phenomenon of soil organisms are suitable for sixth graders' age. Direct experiences with cause–effect relations between abiotic and biotic system elements in combination with extensive reflection about the phenomena and the inquiry activities are expected to promote the pupils' conceptions of validity and their aptitude in specific investigation designs. Thus, we hypothesized that the students would be able to distinguish between good designs and poor designs with regard to sample sizes and observational periods. On the other hand, we expected that the complexity of the ecological lesson content and of the SSI context would lead to weaker learning effects than those found in studies using less demanding domains or learning contexts (e.g. Ehmer, 2008).

We further assumed that the students of the treatment group would experience more autonomy than the participants of the control groups. This is a reasonable assumption since the lesson structure followed a cognitive apprenticeship approach with phases of open-ended experimentation, which is not typical in German science classrooms.

Method

Design

To increase the study's ecological validity, we conducted a field study within the students' usual classroom setting. This, of course, prevents the option of randomly assigning students and teachers to particular study conditions. Thus, in assigning classes to conditions we took care that treatment and control groups were as similar as possible with respect to class size, the ratio of boys to girls, and school location (e.g. middle-class German suburbia). Variables such as general cognitive abilities, levels of competence, family background (which determines verbal ability, socioeconomic, and cultural environment), and institutional conditions such as the quality of the teachers, and support for students outside of the school were not controlled for. Nonetheless, we chose schools that had several classes at the same grade level ('parallel' classes) in order to minimize the number of influencing factors. We chose the teachers in such a way that gender, age, and apparent motivational tendencies were homogeneously distributed among the experimental and control groups.

The research questions were investigated using a quasi-experimental pretest/posttest design. By using a pretest, we tried to statistically control factors (such as the pretest scores and the achievement in several school subjects) that we considered to be relevant for our research questions. All these measures were taken to increase the internal validity of our study and to rule alternative explanations for post-treatment effects between conditions (for a more detailed methodological discussion on this issue, see Shadish, Cook, & Campbell, 2002). In doing so we tried to achieve a suitable compromise between methodological and practical requirements in pursuit of our research goals.

Experimental Conditions

Table 1 shows the study design. In the treatment group (EXP), students received precisely worked-out, competence-oriented lessons on forest ecology by pre-trained teachers. Standard implementation of the EXP treatment was aimed for by providing the participating teachers with a detailed teaching manual that included selected media, methods and learning activities, and through training workshops, given by members of our research group to the teachers and employees of the nature conservation center. We used two control groups: students in the first control group (CG_{ECO}) received the same content as the treatment group, but the teachers were not instructed to follow a particular concept for promoting experimental problem-solving ability. In the second control group (CG_0), the lessons included other science content from the sixth-grade curriculum not in the area of ecology: in southwestern Germany, sciences (biology, chemistry, and physics) in the 'Realschule'-School type are taught as an integrated subject, referred to as *natural science work*. These teachers were also not restricted in their choice of teaching method. The second control group was the result of additional research questions that are not reported in this publication. The time on task was the same in all three groups.

We think that withholding our specific promotion concept from both control groups for the time of running the study is justifiable; first, 'conventional' *natural science work*

Table 1. Experimental conditions (main study)

Aspect		Experimental condition		
		EXP	CG_{ECO}	CG_0
Specific treatment to promote experimental problem-solving ability		Yes	No	No
Forest ecosystem topic		Yes	Yes	No
Partial sample size (*n*)		129	105	106
Age	*M*	11.88	11.84	11.87
	SD	0.58	0.56	0.44

Note: *n*, number of test subjects in the partial sample; *M*, mean value; and *SD*, standard deviation.

lessons in the CG_{ECO} control group had to take into account the enhancement of inquiry competencies, too (KMK, 2005). Second, we delivered the teaching manual to every control group teacher after completing the study. Additionally, we reported the first preliminary results to encourage using the materials in the respective classes later, too.

Sample

Our study comprised 340 students in southwestern Germany (see Table 1). The mean age was 11.9 years (SD = 0.51, range: 11–14 years). The percentage of boys (56%) was slightly higher than that of girls. German is the native language of 80% of the mothers and fathers of the surveyed students. The groups did not differ with respect to the percentage of students with a mother or father who does not speak German as a first language. The students were from urban, respectively, provincial schools, and they are representative for southern Germany. Most of the students' families belong to the economic middle class and have no migratory background. A total of 27 dropouts were due to relocations and illness-related absence; these study participants were not included in the analyses.

We chose sixth-grade students as participants in our study for several reasons: The complex ecosystems domain is anchored in the German sixth-grade *natural science work* curriculum. Its use in the current study can thus be deemed as curricular-valid content. According to the science curriculum, ecological and environmental aspects are dealt with sixth, ninth, and eleventh grades. Why did we choose to implement our study in grade six, although the subject is usually treated in greater depth in the higher grades? In view of Arnold's et al. (2013) findings in higher grades (see earlier), it makes sense to investigate how a spiral curriculum could be implemented to optimally promote students' awareness of appropriate durations of observations and sample sizes. If a fundamental level of awareness for these aspects of higher external and ecological experimental validity could be achieved in a lower grade, in combination with basic cognitive experimental problem-solving competencies, then this would be a solid basis from which to start developing higher levels of this component of experimental problem-solving ability. Various authors have reported on the effectiveness of their interventions enhancing cross-domain experimental skills in grades six or seven, or even lower age groups (e.g. Carey et al., 1989; Ehmer, 2008; Hof, 2011; Klahr & Nigam, 2004). In grade six, most children are expected to be within the transition from the concrete operational phase to the formal operational phase in their cognitive development. This step seems to be an important prerequisite to the further development of cognitive and metacognitive experimental competencies (Schneider et al., 1998) and could possibly facilitate the development of a basic awareness level concerning external validity.

A further comment needs to be made on the student selection. Germany has a three-tier school system. When comparing the results of this study with others conducted in Germany (e.g. Ehmer, 2008; Ganser & Hammann, 2009; Hof, 2011), it is important to notice which kind of schools participated. 'Gymnasien' serve students who are expected to pursue a university degree. Students at these schools generally perform better on tests in all subjects, and in particular in science literacy (Mayer et al., 2008; Prenzel et al., 2013). Students who require more intensive instructional support attend 'Hauptschulen'. We chose to investigate students in 'Realschulen', schools which constitute the middle track of the three tiers, because they tend to perform in the middle of the achievement spectrum. Thus, we tried to reduce the possibility of exceptional student abilities skewing our results or introducing more scatter into the data (due to particularly low- or high-achieving students).

Table 1 shows our attempt to achieve a balanced experimental plan.

Treatment Lessons

We chose 'forest' as the ecosystem to study because of its familiarity to most students in Germany. Our treatment was intended to promote experimental problem-solving ability. It was designed according to moderately constructivist principles. Consistent with Ehmer (2008) and Ganser and Hammann (2009), we designed our teaching unit to be a systematic, cumulative, explicit training, taking into consideration students' understanding, misconceptions, and prior competency levels. The goal was to provide a framework of problem-oriented learning opportunities with varying degrees of direct, teacher-led instruction and open-structured experimentation phases. Through the practical application of scientific inquiry methods and the adherence to natural scientific thinking and working, students were expected to acquire ecological content knowledge, scientific reasoning, and inquiry competencies from experimentation.

A land-use conflict was employed as the background story, representing an authentic context for situated learning. Through this story, students could experience the need for scientific inquiry to investigate ecological causalities and for multi-perspective thinking and responsible decision-making in situations relating to sustainability (Parchmann et al., 2006): Different options for interventions in a forest ecosystem were discussed from ecological, socio-cultural, and economic perspectives in order to find the best option for sustainable development.

The treatment was divided into the following five main phases:

The first lessons introduced important content aspects from the domain of ecology: ecological, economic, and social functions of forests; interactions between organisms and their environment (using the wood grouse as an exemplary endemic bird species); the vertical zoning of forest vegetation; food chains and food webs and the matter cycle in ecosystems. The first encounter of the students with these concepts and terms was contextualized within the background story. The information was repeated, applied, and then presented in greater depth using a problem-oriented structure in the subsequent learning activities.

The second phase began with a one-day visit to a nature conservation center in the *Black Forest* National Park in Southern Germany. Here, the students learned the principle of hypothesis-guided research and experimentation and then put this knowledge into practice in the authentic natural environment of the surrounding forest ('forest classroom'). This place was chosen to also give a real setting for a land-use conflict. It was assumed that by engaging students to conduct their own research projects (cf. Žoldošová & Prokop, 2006a), the experiential process would result in a better understanding of the natural phenomena and the scientific inquiry activities. It was found in a similar learning environment that sixth-grade students' attitudes toward science and ecological knowledge can be promoted by participating in such short-term fieldwork activities (Prokop, Tuncer, & Kvasničák, 2007). We also expected positive effects on social and affective features such as ability to cooperate, peer relationships, and learning motivation as Amos and Reiss (2012) could observe in their study with students of the same age group.

In the third lesson phase, students learned cross-domain aspects of experimentation, including generating epistemic research questions in the pre-experimental phase (which can be advantageous for the planning of experiments; Neber & Anton, 2008), hypothesis testing using data, the need for experimental control, the control-of-variables strategy, the interpretation of evidence, and planning experiments in test series, which they applied then to an ecological context. Our method placed great emphasis on the mental activity of the students, already in the pre-experimental phase (Neber & Anton, 2008). Various examples of experimentation and of sub-contexts in ecology should facilitate the transfer.

In the fourth phase, students learned that experiments in the complex domain of ecology require repeated measurements, long-term observations, and an understanding of the effects of sample size. The students conducted ethological and long-term experiments with regard to ecological issues (e.g. taxis phenomena of woodlice; influences on seed germination and plant growth under different conditions and through different substances). In gender homogeneous learning teams, the pupils reflected on the internal and external ecological validity of their experimental design. Using the cognitive apprenticeship approach, we administered phases with explicit instruction (e.g. showing exactly what the students could do or 'modeling') as well as phases with tutorial support ('scaffolding') in the form of worked examples or process worksheets.

In the fifth phase, the students assumed the roles of science experts. During a second visit to the nature conservation center, they conducted open-ended model experiments on the protective functions of forests in mountainous regions (Figure 1). After the inquiry phase, the learning teams (same as in the fourth phase) presented their experimental procedure and their findings. The rest of the class served as the *science community* to evaluate the quality of the experimental designs and to discuss the aspects of validity and the conclusions of the presenting teams. We assumed that this procedure would facilitate awareness of well-planned, nonconfounded experiments, of the meaning of internal and external ecological validity, and of the limitations of model experiments.

In the final lesson, students assessed the quality of given experiments from fictitious natural scientists. This procedure served to review acquired concepts and strategies of experimentation.

Figure 1. Students plan and conduct their own model experiments, for example, on wind erosion and the function of protective forests

Operationalization of Variables

Operationalization of independent variables. We implemented three experimental conditions, based on the independent variables *didactic-methodical teaching concept* and *subject of the lessons*. The teaching concept was varied by implementing specific measures to promote experimental problem-solving ability or by withholding the measures from the two control groups (CG_{ECO} and CG_0). Additionally, the subject of the lessons was varied by providing information on forest ecosystems to the experimental group (EXP) and to the CG_{ECO} control group but withholding this information from the CG_0 control group.

Operationalization of Dependent Variables

Achievement test. We developed and used a written achievement test based on classical test theory to assess experimental competencies. Compared with performance assessment methods (e.g. Klahr & Nigam, 2004), with interviews about conceptions on experimentation and inquiry strategies (Carey et al., 1989), and with computer-based tools with logfile-based data (e.g. Thillmann et al., 2009), paper-and-pencil tests show a lower level of validity (Germann et al., 1996) because they are not process-related and cannot represent the entire problem-solving process with all its required competencies and their interactions (Hammann et al., 2008). Nevertheless, we used a paper-and-pencil test in view of our big sample and limited resources (cf. Ehmer, 2008).

The test comprised four subtests consisting of 1–8 items (see Table 2):

(1) The *design an experiment* subtest used a partial credit system (0–6 points) to measure the competencies of understanding and considering experimental controls within a two-factorial experiment. This assesses whether students correctly apply the variable-control strategy and compare results with a proper experimental control. The model of Ehmer (2008) and Hammann et al. (2008) was used as the basis for designing the open-response items and coding rules. The evaluation formula was refined to be able to more precisely measure performance differences. Only one item was used for pragmatic reasons (cf. Ehmer, 2008).

(2) The operationalization of the competence *epistemic research questions* was based on criteria stated by Germann et al. (1996), Mayer et al. (2008), and Neber and Anton (2008). Open-response items measured the competence to formulate knowledge-generating and process-regulating questions which address causal links between independent and dependent variables. Our adaptation refers only to the type of formulation and does not account for prior knowledge. This assessment better suits the abilities of sixth graders. The partial credit system ranged from 0 to 4 points.

(3) The *compare approaches* instrument was based on the preliminary work of Ehmer (2008) and Hammann et al. (2008) who called the competence *planning experiments*. Using multiple-choice items, the instrument measured the ability to identify from prespecified approaches the controls suitable for comparison in a two-factorial experiment (control-of-variable strategy). The coding was in accordance with the partial credit model (0–3 points) that assessed aspects of the selection strategy (e.g. 'vary one thing at time' or 'hold one thing at time').

(4) The newly developed *judge the validity limitations of ecological experiments* subtest consisted of multiple-choice items with dichotomized coding (true/false) to measure the application of knowledge regarding the importance of large sample sizes and longer observation periods in experiments in ecological (sub-)systems.

In addition, the matrices subtest from part 1 of the CFT 20-R (Weiss, 2006) was used to measure *basic non-verbal intelligence*.

Table 2. Scale properties (main study)

		Measuring time			
		Pretest		Posttest	
Subtest	Number (items)	n	Cr. α	N	Cr. α
Epistemic research questions	3	331	.54	324	.62
Compare approaches	4	306	.45	309	.59
Judge the validity limitations of ecological experiments	8	280	.52	304	.65
Sense of autonomy	6	234	.64	233	.88

Note: n, scope of the partial sample and Cr. α, Cronbach's alpha.

Other dependent variables. We further asked the participants to report on a 6-point Likert scale the extent of openness, independence, and autonomy (adapted from Kunter, 2005) they experienced in the treatment. The items were also administered to participants in a control group. Additionally, the students in the treatment group described the effects of the learning environment on their attitudes and learning motivation.

Procedure

The sensitivity of the instrument was demonstrated in pilot studies in grades 6 and 9, and showed sufficient item variance. We then improved the treatment and the test method which was also re-tested. The main study was conducted in Spring/Summer 2010. After processing the achievement test and questionnaire, a multi-week period followed with approximately 13 school hours of teaching in all experimental conditions. Classes for the EXP treatment group were additionally held on two consecutive days at the nature conservation center, where employees implemented the program. The posttest took place around two weeks after the last teaching unit.

Data Analysis

We analyzed the data statistically using the statistic software PASW 18 and evaluated in accordance with the *per-fiat* principle using parametric tests (see below).

Results

Test Instruments

As expected, *reliability values* were higher in the posttest than in the pretest (cf. Table 2). The internal consistency of the achievement test subscales was sufficient for group comparisons. The moderate internal consistency of some scales can be attributed to the restriction in the number of items per subtest scale.

The *objectivity* of implementation was ensured through detailed instructions on data collection which was carried out by the participating teachers. We developed and optimized differentiated coding schemes for the objectivity of evaluation in the case of open-response items. The following information on the quality of coding conformity is based in each case on more than 35 double codings. As interval-scale level was assumed, both the mean values of the intraclass correlation ($ICC_{2,1}$) and the chance-adjusted measure of absolute agreement (Cohen's kappa) are specified as measures of agreement below $r = .82$ and Cohen's $\kappa = .63$ (*epistemic issues*); and $r = .98$ and Cohen's $\kappa = .95$ (*design an experiment*). Thus, the intercoder agreement and reliability indicate acceptable to very good values.

In view of the *construct validity*, bivariate correlations among the experimental problem-solving ability variables with school marks (mathematics, natural science work, and German) and general cognitive abilities were calculated. The small coefficients indicate that the dimensions can be differentiated and represent something other than technical achievement within the subjects or general cognitive abilities.

Hypothesis Testing

Effects of the treatment. The MANOVA analysis with univariate post hoc tests showed that the three student groups did not differ with respect to their general cognitive abilities and marks in mathematics. There were differences in relation to marks in German and the integrated subject *natural science work*. These variables were included in covariance analysis as covariates if they contributed significantly to the explained variance in competence development. The MANOVA analyses revealed that the three groups did not differ on the *experimental problem-solving ability* components reported at the time of the pretest, $F(10,598) = 0.96$, $p = .48$.

The posttest values in all groups (apart from several variables in the CG_0 group) were higher in the achievement test than in the pretest (cf. Table 3). These findings suggest a learning effect from working with the test.

To clear up our first research question, we tested the directional hypothesis of differences in learning effects between the experimental groups by ANCOVAs with planned comparisons and the pretest values of the respective dependent variables as covariates. The treatment group was contrasted with the pooled results of the two comparison

Table 3. Descriptive statistics and parameter estimations of the achievement test and one variable of the personality test (main study)

			Measuring time				ANCOVA	
			Pretest		Posttest		Parameter estimation	
Subtest	Group	*n*	*M*	*SD*	*M*	*SD*	M_b	*SE*
Design an experiment	EXP	120	1.56	1.52	2.13	1.84	2.15	0.15
	CG_{Eco}	93	1.58	1.74	1.73	1.79	1.74	0.17
	CG_0	100	1.65	1.59	1.84	1.75	1.82	0.17
Epistemic research questions	EXP	120	1.01	0.67	1.31	0.86	1.31	0.06
	CG_{Eco}	93	1.10	0.64	1.14	0.77	1.17	0.07
	CG_0	100	1.15	0.64	1.18	0.67	1.11	0.07
Compare approaches	EXP	120	1.70	0.70	1.97	0.72	1.97	0.07
	CG_{Eco}	91	1.72	0.62	1.79	0.78	1.78	0.08
	CG_0	98	1.77	0.79	1.70	0.80	1.71	0.08
Judge the validity limitations of ecological experiments	EXP	118	0.45	0.25	0.46	0.27	0.46	0.02
	CG_{Eco}	92	0.39	0.22	0.43	0.24	0.45	0.03
	CG_0	99	0.47	0.24	0.42	0.28	0.40	0.03
Sense of autonomy	EXP	124	2.59	0.90	3.14	1.04	3.09	0.11
	CG_{Aut}	111	2.24	0.97	2.61	1.38	2.68	0.11

Note: *n*, number of test subjects; *M*, mean value; *SD*, standard deviation; M_b, estimated adjusted mean value (corrected); and *SE*, standard error.

groups CG_{Eco} and CG_0. The *pretest value* covariate had a moderate to strong effect on the explained variance for all variables. The hypothesis that the scores of the treatment and the control groups would differ was confirmed for three of the four achievement variables, *design an experiment*: $F(1, 309) = 3.88$, $p < .05$, partial $\eta^2 > .01$; *epistemic research questions*: $F(1, 294) = 4.26$, $p < .05$, partial $\eta^2 > .01$, and *compare approaches*: $F(1, 304) = 6.75$, $p = .01$, partial $\eta^2 = .02$, but could not be confirmed for *judge the validity limitations of ecological experiments*: $F(1, 305) = 1.64$, $p = .20$.

In view of our second research question, the results of the questionnaire show that the treatment allowed the students more autonomy, independence, and openness in conducting experiments on their own compared with the control group participants as indicated from the ANCOVA analysis with the pretest scores as covariates: $F(1, 232) = 6.95$, $p < .01$, partial $\eta^2 = .03$. Although only a moderate effect size could be observed, this result suggests that the cognitive challenges were likely to be greater for the members of the treatment group compared with conventional teaching and learning in *natural science work*. In addition, the students in this group reported that getting to know the job from scientists and doing inquiry on their own were very motivating—as also the visit to the nature conservation center and the fieldwork activities.

Other Findings

Post hoc analyses showed that the significant differences between the treatment group and comparison groups on *design an experiment* and *formulate questions* variables are based on stronger competence development among girls. Girls' performance in *design an experiment* was better in the treatment group compared with the mean of the girls in the control groups—with a small to medium-grade effect size: $F(1, 130) = 4.57$, $p < .05$, partial $\eta^2 > .03$. Within the treatment group, the performance of the girls topped the mean of the boys with a medium-sized effect.

Discussion

In view of our first research question, we established that the treatment was effective in the most of the experimental problem-solving ability dimensions—in contrast to Ganser and Hammann's (2009) findings. In detail, the treatment promoted the competencies of generating epistemic questions, planning two-factorial experiments, and identifying appropriate experimental controls. Furthermore, with regard to our second research question, our moderately constructivist learning environment with phases of open inquiry increased the students' sense of autonomy, as expected.

Given the small number of students and teachers involved, and the use of written paper-and-pencil tests for learning assessments, the effectiveness of the treatment of the EXP group on competence enhancement needs to be viewed with caution. For example, we cannot completely control for factors that affect teaching effectiveness that are innate to the individual teachers involved. As we sampled the middle-grade secondary students in 'Realschulen', the present study could not clarify whether the promotion concept would have been more successful in 'Gymnasien' (cf. Ehmer, 2008; Hof, 2011). We also have no evidence on effects in low-grade secondary schools ('Hauptschulen'). Because we had no other control groups with the specific promotion concept for experimental problem-solving ability within another domain, or in a less challenging context (which presumably causes less cognitive load), our study could not investigate the effectiveness of the treatment per se, respectively, within another domain or learning context.

The retrospectively detected difference in achievement levels between girls and boys is an explorative finding. Thus, our procedure does not allow any interpretation of possible causes concerning gender aspects of the teaching methods (cf., e.g. Goldstein, & Puntambekar, 2004).

Nevertheless, with regard to our study's findings—compared with the results of other investigations—there are grounds for the assumption that basic components of problem-solving ability could probably be acquired more easily in less challenging learning contexts (cf. Ehmer, 2008; Hof, 2011) and less complex domains (e.g. Ehmer, 2008; Ganser, & Hammann, 2009).

Similar to the findings of Prokop et al. (2007) and of Žoldošová and Prokop (2006b), the students of our treatment group reported an increase in their curiosity in several respects (e.g. doing inquiry on their own, gaining an insight into the 'scientists' world', nature conservation as a learning topic). This differs from the results of Žoldošová and Prokop (2006a), who did not find any differences in perceived curiosity (in view of the preferred motivational orientation) under similar experimental conditions amongst Slovakian sixth graders who participated in a field center course—compared with a control group with traditional classroom lessons.

Conclusions and Educational Implications

In view of earlier studies using less complex domains (e.g. Ganser & Hammann, 2009), our study's data show that experimental problem-solving ability can be promoted even in the complex domain of ecology as early as grade six in a challenging, problem-oriented context. We found that average ability sixth-grade science students do not need to be limited to less challenging, isolated contexts or less complex domains that were used in other studies (e.g. Ehmer, 2008; Ganser, & Hammann, 2009), in principle. Our moderately constructivist learning environment is interesting motivating for the students. Moreover, it not only promotes the students' experimental problem-solving ability at least as much as conventional competence-oriented *natural science work* lessons, but also allows more autonomy. Furthermore, the competency to formulate epistemic questions can be stimulated earlier than presumed (Neber & Anton, 2008: 10th-grade students). Nevertheless, the enhancement of experimental problem-solving ability that we find here is small—and smaller than that seen in other studies (e.g. Hof, 2011). This is particularly remarkable in view of the notable treatment effects observed by Ehmer (2008) in her short-term intervention with 'Gymnasium' students. Finally, sixth graders appear to be too young to be made aware of domain-specific aspects of validity and reliability in ecology experiments.

The investigated demanding SSI context covering the four competence areas (*epistemic methods/scientific inquiry, scientific content knowledge*, *decision-making*, and *communication about science*; KMK, 2005) seems to generate lots of cognitive load (cf. Kirschnet et al., 2006) and thus might be consequently challenging for the promotion of basic inquiry skills in experimentation in classes of low or average performance. Therefore, we believe that domains that are less complex than 'ecology' and less-sophisticated SSI contexts would make more sense when introducing basic cross-domain competencies of experimental problem-solving ability more efficiently—especially in grades 1–6; considering the findings from Ganser and Hammann (2009) as well as from our study, we believe that the level of extraneous cognitive load should be reduced as much as necessary for the promotion of experimental problem-solving ability, if that is the center of attention. This is an important thought, for example, with regard to the ambitious goals and complex approaches in education for sustainable development.

Although we have observed that basic awareness for important validity issues in the domain of ecology was not enhanced within this sample, we consider that ecology is relevant for enhancing students' awareness of certain aspects of experimental validity in complex living systems (cf. McBride et al., 2013)—however, in higher grades. Thus, we recommend incorporating this contextualization within a spiral curriculum when students have already acquired fundamental competencies of cognitive experimental inquiry skills. Our study's findings seem to support the assumption that concepts concerning long-term observations and the size of random samples in biological systems are highly complex and abstract (cf. the results of Arnold et al., 2013). Teachers would need to invest more time with instructional efforts to enhance students' competencies in this area, for example, planning more experiments with living animals or working with computer-based simulations in combination with metacognitive reflection and explicit training in methodological criticism.

Finally, on balance, our study's results indicate that the didactic research in the field of promoting specific challenging components of experimental problem-solving ability needs to clarify the optimal method, school grade, and use of appropriate domains and learning contexts with moderate extraneous cognitive load—particularly for average or lower performing learning groups. Additionally, the findings from our intervention suggest the necessity of generating spiral curricula for the promotion of the comprehensive construct of experimental problem-solving ability.

Acknowledgements

This work was supported by the Office of research development and funding of the University of Education Ludwigsburg.

Disclosure Statement

No potential conflict of interest was reported by the authors.

ORCID

Frank Roesch http://orcid.org/0000-0003-1235-5194

References

Amos, R., & Reiss, M. (2012). The benefits of residential fieldwork for school science: Insights from a five-year initiative for inner-city students in the UK. *International Journal of Science Education*, *34*(4), 485–511.

Arnold, J., Kremer, K., & Mayer, J. (2013). Wissenschaftliches Denken beim Experimentieren—Kompetenzdiagnose in der Sekundarstufe II [Experimentation and scientific reasoning—assessment of competencies in the upper high school level]. In D. Krüger, A. Upmeier zu Belzen, P. Schmiemann, A. Möller, & D. Elster (Eds.), *Erkenntnisweg Biologiedidaktik* (Vol. 11, pp. 7–20). Kassel: Universitätsdruckerei. Retrieved from http://www.bcp.fu-berlin.de/biologie/arbeitsgruppen/didaktik/Erkenntnisweg/2012/Arnold.pdf?1362740309

Berkowitz, A. R. (1997). Defining environmental literacy: A call for action. *Bulletin of the Ecological Society of America*, *78*, 170–172.

Carey, S., Evans, R., Honda, M., Jay, E., & Unger, C. (1989). 'An experiment is when you try it and see if it works': A study of grade 7 students' understanding of the construction of scientific knowledge. *International Journal of Science Education*, *11*, special issue, 514–529.

Cutter-Mackenzie, A., & Smith, R. (2003). Ecological literacy: The missing paradigm in environmental education (part one). *Environmental Education Research*, *9*(4), 497–524.

Deci, E. L., & Ryan, R. M. (2000). The 'what' and 'Why' of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268.

Ehmer, M. (2008). *Förderung von kognitiven Fähigkeiten beim Experimentieren im Biologieunterricht der 6. Klasse. Eine Untersuchung zur Wirksamkeit von methodischem, epistemologischem und negativem Wissen* [Promoting cognitive abilities with experimentation in grade six biology teaching. An investigation of the effectiveness of methodological, epistemological, and negative knowledge] (Doctoral dissertation). University of Kiel. Retrieved from http://eldiss.uni-kiel.de/macau/receive/dissertation_diss_00003034

Ganser, M., & Hammann, M. (2009). Teaching competencies in biological experimentation. In M. Hammann, K. Boersma, & A. J. Waarlo (Eds.), *The nature of research in biological education: Old and new perspectives on theoretical and methodological issues: Proceedings of the VIIth conference of European researchers in didactics of biology (ERIDOB)* (pp. 377–394). Utrecht: Freudenthal Institute for Science and Mathematics Education, Utrecht University.

Germann, P. J., Aram, R., Odom, A. L., & Burke, G. (1996). Student performance on asking questions, identifying variables, and formulating hypotheses. *School Science and Mathematics*, *96*(4), 192–201.

Goldstein, J., & Puntambekar, S. (2004). The brink of change: Gender in technology-rich collaborative learning environments. *Journal of Science Education and Technology*, *13*(4), 505–522.

Hammann, M., Phan, T. T. H., Ehmer, M., & Grimm, T. (2008). Assessing pupils' skills in experimentation. *Journal of Biological Education*, *42*(2), 66–72.

Hof, S. (2011). *Wissenschaftsmethodischer Kompetenzerwerb durch Forschendes Lernen. Entwicklung und Evaluation einer Interventionsstudie* [Science methodical acquisition of competency by inquiry based learning. Development and evaluation of an intervention study] (Doctoral dissertation). Universität Kassel. Kassel: Kassel University Press.

Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundation for the twenty-first century. *Science Education*, *88*, 28–54.

Keselman, A. (2003). Supporting inquiry learning by promoting normative understanding of multivariable causality. *Journal of Research in Science Teaching*, *40*(9), 898–921.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75–86.

Klahr, D. (2000). *Exploring science. The cognition and development of discovery processes*. Cambridge, MA: MIT Press.

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction. Effects of direct instruction and discovery learning. *Psychological Science*, *15*(10), 661–667.

KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (Ed.). (2005). *Beschlüsse der Kultusministerkonferenz. Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004* [Resolutions from the National Conference of Ministers of Education. Education standards in biology for the fifth- to tenth-grade students. Resolution from 2004-12-16]. München, Neuwied: Wolters Kluwer.

Kunter, M. (2005). *Multiple Ziele im Mathematikunterricht* [Multiple objectives in mathematics teaching] (Doctoral dissertation). Münster: Waxmann.

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

Mayer, J., Grube, C., & Möller, A. (2008). Kompetenzmodell naturwissenschaftlicher Erkenntnisgewinnung [Competence model of scientific inquiry]. In U. Harms & A. Sandmann (Eds.), *Ausbildung und Professionalisierung von Lehrkräften. Internationale Tagung der Fachsektion Didaktik der Biologie im VBiO, Essen 2007* [Education and professionalization of teachers. International conference of the section 'Biology Education' within VBiO, Essen 2007] (pp. 63–79). Innsbruck: Studien Verlag.

McBride, B. B., Brewer, C. A., Berkowitz, A. R., & Borrie, W. T. (2013). Environmental literacy, ecological literacy, ecoliteracy: What do we mean and how did we get here? *Ecosphere*, *4*(5), art. 67. Retrieved from http://dx.doi.org/10.1890/ES13-00075.1

Morrone, M., Mancl, K., & Carr, K. (2001). Development of a metric to test group differences in ecological knowledge as one component of environmental literacy. *The Journal of Environmental Education*, *32*(4), 33–42.

Muckenfuß, H. (1995). *Lernen im sinnstiftenden Kontext. Entwurf einer zeitgemäßen Didaktik des Physikunterrichts* [Learning in meaningful contexts. Draft of contemporary didactics of physics education]. Berlin: Cornelsen.

Neber, H., & Anton, M. A. (2008). Promoting pre-experimental activities in high-school chemistry: Focusing on the role of students' epistemic questions. *International Journal of Science Education*, *30*(13), 1801–1821.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Orr, D. W. (1992). *Ecological literacy: Education and the transition to a postmodern world*. Albany: State University of New York Press.

Parchmann, I., Graesel, C., Baer, A., Nentwig, P., Demuth, R., Ralle, B., & ChiK Project Group. (2006). 'Chemie im Kontext': A symbiotic implementation of a context-based teaching and learning approach. *International Journal of Science Education*, *28*(9), 1041–1062.

Prenzel, M., Sälzer, C., Klieme, E., & Köller, O. (Eds.). (2013). *PISA 2012. Fortschritte und Herausforderungen in Deutschland* [PISA 2012. Progress and challenges in Germany]. Münster: Waxmann. Retrieved from http://www.pisa.tum.de/fileadmin/w00bgi/www/Berichtband_und_Zusammenfassung_2012/PISA_EBook_ISBN3001.pdf

Prokop, P., Tuncer, G., & Kvasničák, R. (2007). Short-term effects of field programme on students' knowledge and attitude toward biology: A Slovak experience. *Journal of Science Education and Technology*, *16*(3), 247–255.

Riess, W., & Mischo, C. (2010). Promoting systems thinking through biology lessons. *International Journal of Science Education*, *32*(6), 705–725.

Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching*, *41*(5), 513–536.

Schneider, W., Bullock, M., & Sodian, B. (1998). Die Entwicklung des Denkens und der Intelligenzunterschiede zwischen Kindern [The development of thinking and differences in intelligence amongst children]. In F. E. Weinert (Ed.), *Entwicklung im Kindesalter* [Child development] (pp. 53–74). Weinheim: Beltz Psychologie-Verlags-Union.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton-Mifflin.

Smith, T. M., & Smith, R. L. (2009). *Ökologie* [Ecology]. (6th ed.). München: Pearson Studium.

Thillmann, H., Künsting, J., Wirth, J., & Leutner, D. (2009). Is it merely a question of 'what' to prompt or also 'when' to prompt? The role of point of presentation time of prompts in self-regulated learning. *Zeitschrift für Pädagogische Psychologie*, *23*(2), 105–115.

Weiss, R. H. (2006). *Grundintelligenztest Skala 2* [Scale 2 basic intelligence test] *(CFT 20-R)*. Göttingen: Hogrefe.

Žoldošová, K., & Prokop, P. (2006a). Analysis of motivational orientations in science education. *International Journal of Science and Mathematics Education*, *4*(4), 669–688.

Žoldošová, K., & Prokop, P. (2006b). Education in the field influences children's ideas and interest toward science. *Journal of Science Education and Technology*, *15*(3), 304–313.