International Journal of Innovation in Science and Mathematics Education, 29(2), 19-31, 2021

Fostering Learning with Incremental Scaffolds During Chemical Experimentation: A Study on Junior High School Students Working in Peer-Groups

Mirlinda Mustafa[a*], Anestis Ioannidis[a*], Laura Ferreira González[a*], Timo Dabrowski[b], Jörg Großschedl[a*]

Corresponding author: Jörg Großschedl (j.grossschedl@uni-koeln.de)
[a] Institute for Biology Education, University of Cologne, Köln 50931, Germany
[b] Dreikönigsgymnasium, Köln 50931, Germany
*These authors contributed equally to this article.

Keywords: conceptual knowledge, incremental scaffolds, inquiry-based learning, junior high school students

Abstract

Scaffolds are considered to be a promising method of supporting learning. In this study, we investigated the learning efficacy of scaffolds in an inquiry-based learning scenario. Three tasks posed a question/problem to facilitate inquiry-based learning, and scaffolds offered the answer/solution in multiple steps (so-called incremental scaffolds). The use of the scaffolds was voluntary and students' learning efficacy was compared with a traditional teaching approach. A total of $N = 105$ seventh graders participated in the quasi-experimental study. Incremental scaffolds were available to the students in the treatment group. Students in the control group received the same question/problem but could only ask the teacher about the answer/solution. Concept maps were used at pre- and posttest to assess conceptual knowledge acquisition. In-line with our hypothesis, results show that students in the treatment group outperformed controls concerning conceptual knowledge acquisition. Regarding the number of misconceptions students used, there were no differences between the groups. Our study indicates that incremental scaffolds are an appropriate method to provide students with the exact help they really need. Based on our findings, we offer practical implications and recommendations for future research.

Introduction

Inquiry-based learning

Inquiry-based learning is a pedagogical approach for student-centred learning, which starts by posing questions, scenarios, or problems. Learners address these issues through inquiry and intellectual engagement in order to find answers/solutions, to develop a deeper understanding of the underlying concepts, and to acquire conceptual knowledge (Hmelo-Silver, Duncan, & Chinn, 2007; Tan, Koppi, & Field, 2016; Wangdi, Precharattana, & Kanthang, 2020). Inquiry-based learning improves the students' independence, as they undergo the research process as a whole (Mieg, 2019), from developing questions and hypotheses, selecting the methods, and presenting the results (Pedaste, Mäeots, Leijen, & Sarapuu, 2012). Inquiry-based learning is cognitively demanding, requires advanced metacognitive skills and a high degree of motivation (Thomas, Bennett, & Lockyer, 2016; Zimmermann, 1998). Students facing these challenges often need instructional support (Schmidt-Weigand, Hänze, & Wodzinski, 2009). Tight instructional support (i.e., guidance) decreases students' motivation by restricting their scope of action, but too little instructional support also negatively affects motivation because of a high risk of failing the task (Bjonness & Kolsto 2015; van de Pol & Elbers, 2013). Therefore, scholars recommend a middle ground of instructional support, which gives the students the responsibility for the learning process (Haltunen, 2003) and adapts the teacher's degree of control in such a way that guidance is tailored to students' needs. Providing an adequate level of instructional support is particularly challenging in an increasingly heterogeneous classroom (Forghani-Arani, Cerna, & Bannon, 2019). Hence, adequate instructional support which facilitates step-by-step learning, is of great importance (Hmelo-Silver et al., 2007).

Incremental scaffolds

Scaffolds provide a solution or answer to a problem or question that arise from a particular task. Sometimes, scaffolds also give additional information or include prompts (i.e., hints to find the answer/solution) to support the students. A particular type of scaffolds, so-called incremental scaffolds, presents the answer/solution step-by-step to the students (Schmidt-Wiegand, Franke-Braun, & Hänze, 2008) or gradually provides additional information or prompts. The step-by-step presentation reduces the complexity of the task with the aim to decrease the necessary amount of working memory resources (so-called cognitive load; cf. Sweller, van Marrienboer, & Paas, 1998). Scaffolding enables a learner "to solve a problem, carry out a task or achieve a goal which would be beyond his unassisted efforts" (Wood, Bruner, & Ross, 1976, p. 90). Franke-Braun, Schmidt-Weigand, Stäudel, and Wodzinski (2008) emphasise that incremental scaffolds are particularly useful for students requiring special assistance. Incremental scaffolds provide the support these students actually need without exposing their lack of knowledge in the classroom.

Schmidt-Weigand et al. (2009) compared three ways of supporting students' problem solving in a collaborative context: The solution was given (1) at once (worked-out examples), (2) in multiple steps without (incremental scaffolds) or (3) with particular prompts like questions, graphics, or hints to promote active thinking (incremental strategic scaffolds). They found that an incremental scaffolding increased the motivation of the students, improved their feeling of competence, and led to more problem- and regulation-focused communication between students working in pairs. This is supported by recent research, highlighting the positive effects of scaffolding on performance and motivation (Krause, Stark, & Mandl, 2004; Lou, Abrami, & D'Apollonia, 2001).

A source of scaffolding can be 'soft' or 'hard' (Saye & Brush, 2002). Soft scaffolding refers to the support provided by a teacher or a peer when required. It demands constant monitoring of students' performance to provide the right amount of help at the right time. Due to practical difficulties with this concept in a class of 20-30 students, hard scaffolding plays an important role. Hard scaffolds are paper-and-pencil or electronic tools which anticipate the particular needs of the students when learning. They often consist of a question, a hint, or a prompt that stimulates students to think in more depth about a question or problem (Belland, Glazewski, & Richardson, 2008). However, we are not aware of any research that compares learning with incremental scaffolds (hard scaffolds) with instructional support from the teacher (soft scaffolds; Arnold, Kremer, & Mayer, 2017). As the latter is the most common way of classroom learning, this comparison has high ecological validity.

Conceptual knowledge

To be knowledgeable in science implies the understanding of how scientific concepts are interrelated. These interrelated concepts are stored in long-term memory (Veríssimo Catarreira, Godinho Lopes, Casas García, & Luengo González, 2017) and represent a person's conceptual knowledge (Krathwohl, 2002). Conceptual knowledge comprises the knowledge of facts, laws, and principles, all of which are necessary for dealing with different tasks in the academic environment (Furtak, Seidel, Iverson, & Briggs, 2012). Conceptual knowledge makes it possible to edit and understand multiple stimuli, such as words, sounds, or pictures and allows the expression of knowledge in both verbal and non-verbal ways. It also mediates the generalisation or transfer of knowledge from one domain to another (Lambon-Ralph, Pobric, & Jefferies, 2009).

Research question

Schmidt-Weigand et al. (2008, p. 373) show that students supported by incremental scaffolds when learning the scientific concept of density, outperformed the control group in a paper and pencil test measuring students' understanding of density (medium effect of Cohens $d = 0.49$; $N = 63$ ninth graders). Contrarily, the research of Franke-Braun et al. (2008) shows only a minor benefit of incremental scaffolds; they explain these results with the complexity of the incremental scaffolds, the lack of students' experience with learning in pairs as well as the lack of distinct metacognitive abilities ($N = 62$ ninth graders). However, it is difficult to compare those findings as the studies differ in regard to the research foci, the control groups used, and the learning approaches. Consequently, our research question is whether an incremental scaffolding (step-by-step; treatment group) leads to a significant increase of conceptual knowledge in comparison to a traditional approach, where students seek instructional support from the teacher if they have questions (control group).

Methods

Sample

We conducted an a priori power analysis with G*Power software (Faul, Erdfelder, Lang, & Buchner, 2007) to determine the required sample size using the following specifications: we considered a medium effect to be desirable for our treatment ($f = .25$; cf. Hattie [2012, p. 15]), defined an α error level of .05, and a power of $1 - \beta = .90$ (according to Whitley and Ball [2002], the power represents the chance of correctly identifying a significant difference between two groups when the difference really exists in the population). Based on these specifications, the power analysis revealed a required sample size of 128 students. Our actual sample consisted of 132 students ($n = 65$ male, $n = 67$ female) from an integrative comprehensive secondary school in Germany (five seventh-grade classes) aged between 12 and 14 years old ($M = 12.43$ years, $SD = 0.63$). Due to a number of dropouts, only the data of $N = 105$ students could be included in our analysis. Despite this relatively high drop-out rate of 20%, a post-hoc power analysis (Faul et al., 2007) was conducted which still showed an adequate power of $1 - \beta = .81$ for the reduced sample size. A total of $n = 59$ students (50.8% male; $M = 12.42$ years, $SD = 0.59$) attended the treatment group, $n = 46$ students (47.8% male; $M = 12.39$ years, $SD = 0.68$) attended the control group.

Research design and procedure

To address our research question, we conducted a quasi-experimental intervention study with a pre- and posttest design. The pretest, the intervention, and the posttest were all carried out on consecutive days, and both tests involved the same concept mapping-task to assess conceptual knowledge. Concept maps are diagrams with a network structure consisting of nodes (concepts) which are connected via labelled arrows in a meaningful way (Novak & Cañas, 2008). As students are often unfamiliar with the concept mapping-technique, a training session preceded the pretest to prevent inadequate concept mapping-skills by the students. In the intervention phase, which followed the pretest, chemistry classes were randomly assigned to the treatment and control groups. A random assignment of the students to both groups was not possible because of organisational reasons (e.g., restricted availability of parallel laboratory rooms).

Both groups conducted the same two experiments on chemical reactions in pairs and had to work on the same three tasks. The tasks were presented on worksheets, related to the experiments, and posed a question/problem to facilitate inquiry-based learning. Working on the tasks was supported by incremental scaffolds in the treatment group and by the teacher in the control group (see intervention section for more detail). As “learning science means learning to do science” (Brewer & Smith, 2009, p. 14), inquiry-based learning provides “the opportunity to generate scientific knowledge through research” (The President's Council of Advisors on Science and Technology, 2012, p. 25).

Our study strictly adhered to the ethical guidelines of the Declaration of Helsinki and was permitted by our Ministry of School and Further Education (10-45 No. 2). Admission was subject to the condition that participants were informed about (1) aims and process of the investigation, (2) entire voluntariness of participation, (3) possibility of dropping out of participation at any time, (4) guaranteed protection of data privacy (collection of only anonymised data), (5) possibility of requesting data cancelation, (6) no-risk character of study participation, and (6) contact information in case of any questions or problems. Additionally, we obtained the written and informed consent of all participants, as well as of their parents prior to the study.

Concept map-training

The concept mapping-training lasted 60 minutes and the students worked in pairs. Two posters, each showing a concept map on the same topic (‘Life in the wilderness’) were presented to the students on the blackboard. One concept map was correct, the other one contained common mistakes (e.g., arrows pointing in the wrong direction). Students compared both maps and identified the main characteristics of an accurate concept map. After this, they were given a sheet of paper with a pre-constructed concept map on a different topic (‘Teaching natural sciences in school’), with empty nodes and arrows to be filled in with the right words from a given list. The training ended with the students constructing a concept map by themselves on the topic of ‘zoology’ by using a given list of concepts and linking words. This task was similar to the pre- and posttest. To ensure that students could master this task, concept maps were evaluated by the instructor after the training session; a 15-minute feedback was then provided prior to implementing the pretest. None of the example-topics were linked to the content of the test- and learning phases.

Pretest

Students’ conceptual knowledge was assessed by a concept mapping-task. Students got 35 minutes of time to construct a concept map on a focus question on a piece of paper which referred to the target-topic ‘chemical reactions’. A list of 14 concepts (e.g., activation energy, reactant) and 13 linking words (e.g., heat) were provided to the students, no collaboration among students was allowed. Although, a variety of other methods, for example, multiple choice tests (DiBattista & Kurzawa, 2011), interpretive essays (Bolte, 1999), or similarity judgements tests (Großschedl & Harms, 2013) promise a valid assessment of conceptual knowledge, concept maps seemed promising for several reasons. Concept maps start with an open-ended question (the focus question) and provide richer insights into students’ thinking processes and understanding than responses to closed-ended questions like multiple choice questions or the questions of a similarity judgements test. Like interpretive essays, concept maps reveal students’ misconceptions, but their scoring/evaluation requires significantly less effort than in the case of essays (Brandstädter, Harms, & Großschedl, 2012; Großschedl, Mahler, & Harms, 2018).

Intervention phase

The intervention phase lasted 60 minutes and took place in a chemistry laboratory of the school. All students attended a lecture given by the instructor on the topic ‘endothermic chemical reaction.’ In the following 15 minutes, paired students conducted two experiments with copper sulphate. They were then given 20 minutes to work individually on three tasks referring to the conducted experiments. These tasks were identical for the treatment and control groups. In the treatment group, four sets of incremental scaffolds were provided (their use was explained to the treatment group before the intervention phase), addressing each of the three tasks: two scaffold-sets on chemical equations and reactants/products for the first task; one scaffold-set on exothermic/endothermic reactions for the second task, and one scaffold-set on activation energy for the last task. Each incremental scaffold-set consisted of four (A, B, C, D) parts (see Figure 1 for an example; part A of scaffold-set 1) and each part was printed on a sheet of paper and folded three times. By unfolding the sheet of paper, students were given the particular parts step-by-step. These parts included additional prompts or hints, an extra piece of information, or an example from everyday life. The use of incremental scaffolds was optional. The control group received a hint equivalent to those on the incremental scaffolds in the form of a simple solution given by the teacher on request. In case of content-related questions, the students in the treatment group were encouraged to use the scaffolds, whereas the control group received an answer from the teacher when asked. No further content-related questions were answered in any group, whereas general questions (e.g., questions about the time still available) were answered in both groups. The learning phase finished with a presentation of the solution to the entire class, so that all students could correct their answers.

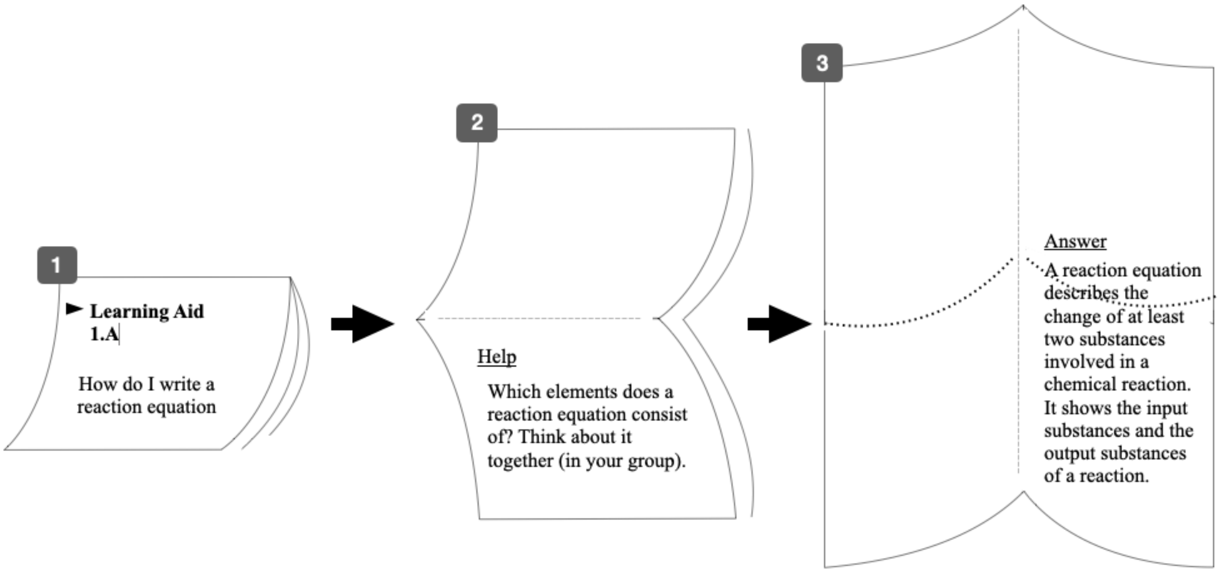

Figure 1. Example of an incremental scaffold from the student´s point of view. Step 1: Number of the incremental scaffold (upper corner) and question. Step 2: After the first unfold, the help is revealed. Step 3: After the second unfold, the answer appears.

Posttest

The posttest lasted 30 minutes and was identical to the pretest.

Analysis of pretest and posttest concept maps

Students’ conceptual knowledge was assessed by a quantitative and qualitative approach. The quantitative approach refers to the relational scoring method developed by McClure, Sonak, and Suen (1999). This method involves the scoring of the individual propositions of a concept map. Propositions consist of two nodes (concepts) connected to each other with a labelled arrow. A completely mistaken proposition, which depicts a relationship between two concepts that is not reasonable from the scientific point of view, was scored with a zero. If two related concepts were connected with each other, but with the wrong kind of relationship (mistaken linking word), one point was granted. In cases where the type of relationship was correct, but the direction of the arrow was wrong, two points were given. An error-free proposition was given three points (for a detailed description of the coding scheme see Figure 2). Students' conceptual knowledge was then expressed as a sum score (so-called proposition accuracy score), which is derived from the summation of the individual sub-scores for each proposition. In order to check the quality of scoring, ten percent of the concept maps from pre- and posttest were randomly chosen and independently scored by a second rater (Döring & Bortz, 2016). Spearman correlation was calculated to determine interrater reliability and indicates reliable scoring ($r = .92$, $p < .001$).

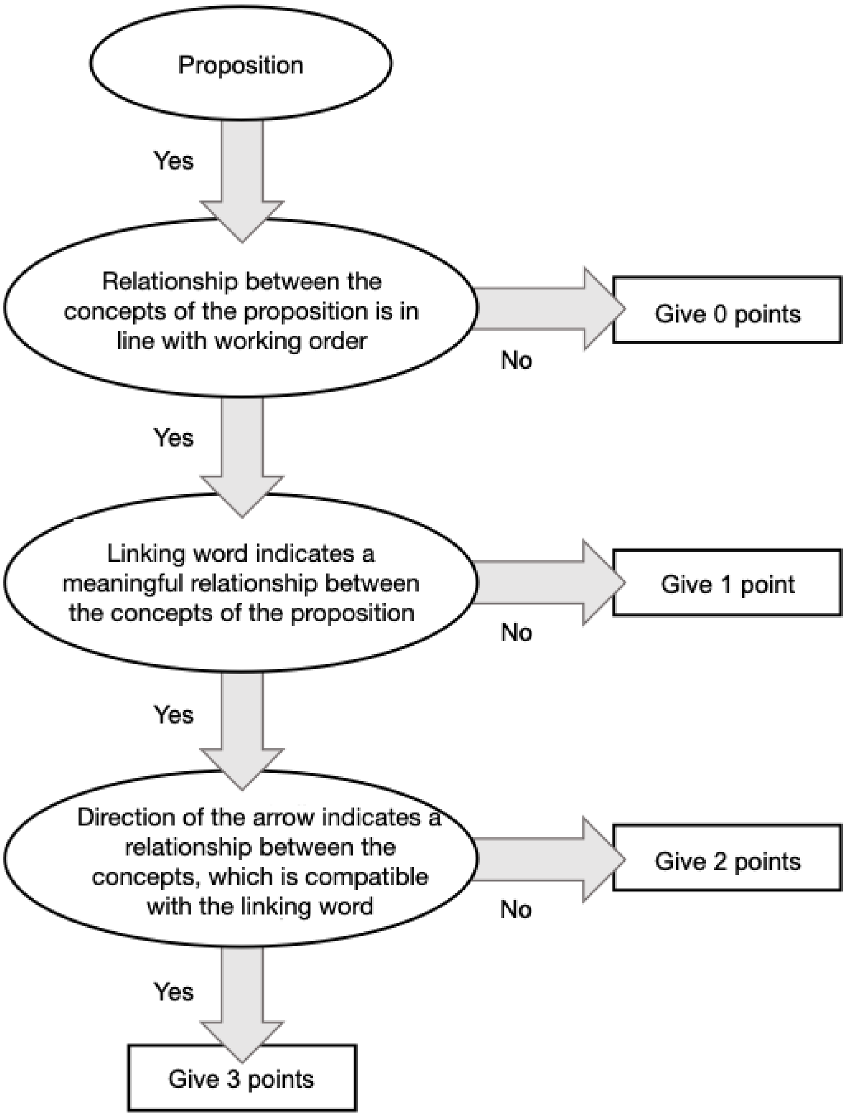

Figure 2. Coding scheme according to McClure et al. (1999; adapted by Brandstädter et al., 2012)

The qualitative approach applies to the assessment of misconceptions. Misconceptions can be defined as ‘any conceptual idea that differs from the commonly accepted scientific consensus’ (Garnett & Treagust, 1990, p. 147). Students already have ingrained ideas and concepts that are inconsistent with, or even in strong contrast to, scientific views (Duit & Treagust, 2003). Therefore, misconceptions pose a challenge for education in science because they can be widespread amongst students and can be resistant to change (Smith, III, diSessa, & Roschelle, 1993). Previous studies show that from middle school to university, students’ knowledge of chemical reactions is limited and characterised by numerous misconceptions (Ahtee & Varjola, 1998). Concept maps can be used to uncover students’ misconceptions (Djanette & Fouad, 2014). Following the qualitative approach of Djanette and Fouad (2014), the analysis of misconceptions involves three steps: (1) identification of all technically wrong propositions, (2) inductive identification of common misconceptions which become apparent in these propositions (see Appendix A), and (3) calculation of sum scores representing the number of misconceptions used (so-called misconceptions score). In order to determine interrater reliability, ten percent of the concept maps from pre- and posttest were also randomly chosen and independently scored by a second rater (Döring & Bortz, 2016). Spearman correlation was calculated which indicates acceptable interrater reliability ($r = .74$, $p < .001$).

Data analysis

We used *SPSS 23* and specified an α level of .05 for statistical analyses.

Results

According to our research question, we were interested in whether an incremental scaffolding (treatment group) improves conceptual knowledge acquisition in comparison to a traditional approach (control group). We used two types of scores as indicators of students’ conceptual knowledge: (1) The ‘proposition accuracy score’ refers to the relational scoring method developed by McClure et al. (1999). (2) The ‘misconceptions score’ represents the number of misconceptions and was suggested by Djanette and Fouad (2014). Conceptual knowledge acquisition should be associated with an increase of ‘proposition accuracy scores’ and a decrease of ‘misconceptions scores’ from pre- to posttest. Before testing these hypotheses, a Shapiro-Wilk normality test was performed which indicated non-normality of all variables. However, histograms showed approximately normal distribution of the ‘proposition accuracy scores’. Considering the sample size (cf. central limit theorem described by Field [2013]) and the histograms, we applied a t-test and repeated measures analysis of variance to the ‘proposition accuracy scores’. In contrast to the ‘proposition accuracy scores’, the ‘misconceptions scores’ showed strong right-skewed distributions in the histograms. According to Bortz and Lienert (2008), we implemented a Solomon four-group design and performed non-parametric Mann-Whitney U tests. No values deviating more than 3 *SD* from the mean were detected in either group. As students were not randomly assigned to the treatment and control groups, we checked whether groups differed in their conceptual knowledge at the pretest which would limit the comparability of the two groups. A t-test showed that the treatment group ($M = 8.95$, $SD = 4.83$) and the control group ($M = 8.93$, $SD = 5.26$; see Table 1) had comparable ‘proposition accuracy score’ in the pretest, $t(103) = 0.02$, $p = .988$. The same was the case for the “misconceptions score’ with comparable mean ranks in the treatment ($M_{\text{Rank}} = 52.17$, $n = 59$) and control groups ($M_{\text{Rank}} = 54.07$, $n = 46$), $U = 1308.00$, $p = .731$ (Mann-Whitney U test).

Table 1. Pre- and posttest mean scores (proposition accuracy scores) and standard deviations as a function of instruction condition (group)

	Pretest		Posttest		
Group	*M*	*SD*	*M*	*SD*	*n*
Treatment	8.95	4.83	15.58	8.88	59
Control	8.93	5.29	12.43	5.97	46
Total	8.94	5.01	14.20	7.86	105

A repeated measures analysis of variance was executed to investigate whether 'proposition accuracy scores' increased from pre- to posttest (within-subject factor 'time') and whether this increase differs between the two groups (interaction effect between 'time' and 'group'). Students' 'proposition accuracy scores' significantly improved in both groups from pre- to posttest, $F(1, 103) = 56.87$, $p < .001$, partial $\eta^2 = 0.356$ (large effect; Richardson, 2011, p. 142). The treatment group achieved a higher increase than the control group, $F(1, 103) = 5.42$, $p < .05$, partial $\eta^2 = 0.05$ (medium effect; Richardson, 2011, p. 142). These results support our hypothesis that conceptual knowledge acquisition benefits from a step-by-step support through incremental scaffolds in comparison to a traditional approach, in which students had the opportunity to ask questions to the teacher.

Due to the non-normality of the 'misconceptions scores', the analysis followed a Solomon four-group design (Bortz & Lienert, 2008). Difference scores between the 'misconceptions scores' from pre- and posttest were calculated to describe the rate of change in the treatment and control groups. These scores were used as the dependent variable in a Mann-Whitney U test with group as independent variable. The Mann-Whitney U test showed no significant difference between the treatment group ($M_{\text{Rank}} = 54.23$, $n = 59$) and the control group ($M_{\text{Rank}} = 51.42$, $n = 46$; $U = 1284.50$, $p = .623$), indicating that incremental scaffolds did not reduce misconceptions more effectively than the traditional approach.

An explorative analysis showed that students had a variety of misconceptions regarding the target-topic 'chemical reactions'. These misconceptions arose in both groups and included the idea that...

- a "chemical reaction" cannot take place without a "gas burner" (e.g., "chemical reaction needs a gas burner").
- a "chemical reaction" is "energy" (e.g., "chemical reactions consist of energy").
- "water" generates "energy" (e.g., "water provides energy").

Beyond that, students incorrectly combined the concepts "energy", "heat", and "activation energy" (examples: "heat generates energy", "activation energy generates heat") as well as "exothermic", "endothermic", and "heat" (examples: "endothermic is exothermic", "endothermic emits heat", "exothermic absorbs heat")

Discussion

This study investigated whether an incremental scaffolding improves conceptual knowledge acquisition in comparison to a traditional approach, where students require instructional support from the teacher if they have questions. Both approaches were implemented in a chemistry classroom following the concept of inquiry-based learning. Concept maps from the pre- and posttest provided insight into the students' conceptual knowledge and were evaluated according to the relational scoring method developed by McClure et al. (1999). Scores that emerged from this evaluation indicate the average accuracy of propositions and their number (so-called proposition accuracy score). Beyond that, the number of misconceptions was determined as suggested by Djanette and Fouad (2014; so-called misconceptions score) to investigate whether incremental scaffolds help to overcome scientifically wrong conceptions (i.e., misconceptions). Both scores allow different statements about the learning efficacy of incremental scaffolds. Whereas the ‘proposition accuracy score’ supports the learning efficacy of incremental scaffolds, the ‘misconceptions score’ does not. At a first glance, this result appears contradictory, but it can be explained by the nature of learning opportunities offered to the students. Whereas these learning opportunities convey the scientific concepts, they neglect students’ misconceptions. Thus, the learning opportunities of this study enable the acquisition of conceptual knowledge, but the misconceptions of the students remain unchanged due to a lack of learning opportunities that encourage cognitive conflict (cf. Posner, Strike, Hewson, & Gertzog, 1982). It is important to note that the standard deviation of the ‘proposition accuracy scores’ (cf. posttest results of the treatment group in Table 1) noticeably exceeds the standard deviation of the control group. This could indicate that incremental scaffolds are only suitable for certain students. Since the scores in the treatment group are distributed upwards more strongly than downwards compared to the control group, we assume that not all students used the scaffolds correctly or benefited from their use.

Only a few studies examined the effect of incremental scaffolds on learning (van de Pol, Volman, & Beishuizen, 2010). Since they use different strategies in the control group (e.g., worked-out examples and incremental strategic scaffolds in the study of Schmidt-Weigand et al. [2009]), it is very challenging to assess the learning efficacy of incremental scaffolds in the science classroom. As the comparison of incremental scaffolds to a traditional approach of teaching and learning has high ecological validity, our study provides further evidence for the learning efficacy of incremental scaffolds.

Beyond the learning efficacy of incremental scaffolds, an explorative evaluation of the concept maps revealed a series of misconceptions about chemical reactions. One misconception concerns the origin of energy and can be described as energy generation (e.g., “heat generated energy”). This misconception contradicts the concept of energy conversions and is being used in various contexts (cf. Barke, 2006; Opitz, Blankenstein, & Harms, 2016). An additional misconception refers to the relationship between the concepts of activation energy, energy, and heat. Although the students noticed a relation between these concepts, they were not able to describe this relation correctly. As a final point, students had problems distinguishing between endothermic and exothermic reactions (e.g., “exothermic is endothermic”), which can be found in other studies too (e.g., de Vos & Verdonk, 1986; Kind, 2004).

Although, the comparison of incremental scaffolds with a traditional approach of teaching and learning has high ecological validity, our study has some limitations:

- The quasi-experimental design of the study restricts its internal validity. However, both groups entered the study with similar conceptual knowledge, suggesting that the conclusions drawn from our study are robust.
- The students were informed about the aim of our study in advance. This may have led to an increased effort of the students (due to extrinsic motivation to impress) and thus might have influenced the results (Rosenthal effect; Rosenthal & Fode, 1963).
- Incremental scaffolds were new to the students. Thus, a novelty effect could have positively influenced the learning efficacy of incremental scaffolds (cf. Kormi-Nouri, Nilsson, & Ohta, 2005).

- Due to a drop-out rate of 20%, we did not achieve the desired power of 1 – β = .90. However, the power is considered sufficient in the range of 80% to 95% (Whitley & Ball, 2002), as is the case for this study.
- Since our study provided a voluntary use of the incremental scaffold for the students, we collected no data about how many students in the treatment group did make use of it. Furthermore, we cannot be sure whether the students used the incremental scaffold in the intended form, or if they just sought for the solution, given on the last card, too early.

In order to gain more valid statements about the learning efficacy of incremental scaffolds further studies are necessary. These studies should record the actual use of the scaffolds by the students (e.g., by camera observations). Moreover, they should apply incremental scaffolds in various subjects (e.g., in chemistry, physics, or mathematics), groups (e.g., students with and without special need), or social settings (e.g., individual work, partner work, group work) to explore the conditions which are suitable for successful scaffolding. Van de Pol, Volman, Ooort, and Beishuizen (2015), for example, found that the learning efficacy of incremental scaffolds depends on how much time students have to complete a particular task. The more time they have, the more scaffolds they use. Furthermore, future studies should investigate the use of incremental scaffolds in everyday school life. The long-term embedding of incremental scaffolds reduces the risk of students being unfamiliar with scaffolds. This in turn prevents a novelty effect that could skew the results of future studies.

It appears noteworthy that incremental scaffolds are not typically utilised in German schools and only a small number of studies have explored the influence of this strategy in inquiry-based learning scenarios. There are various possible reasons why scaffolds are rarely used in the classroom. Some teachers may not be familiar with incremental scaffolds and others may pull back from the effort creating scaffolds. This is suboptimal as incremental scaffolding seems to be a powerful tool in classroom teaching. Giving students the right amount of instructional support is an important challenge in school. Incremental scaffolding gives the students the opportunity to decide for themselves when they need help. This way they can work independently and in a self-regulated manner. Incremental scaffolds can be seen as an efficient way to give the students the help they really need. They also meet the preference of the students to get anonymous help rather than asking the teacher (Franke-Braun et al., 2008). Therefore, our study could serve as a useful example of how incremental scaffolds could be used to support teaching science.

Acknowledgements

We would like to thank Naemi Welter and Britta Mörs for their comments on an earlier version of the manuscript. The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

Ahtee, M., & Varjola, I. (1998). Students' understanding of chemical reaction. *International Journal of Science Education, 20*(3), 305–316.

Arnold, J., Kremer, K., & Mayer, J. (2017). Scaffolding beim Forschenden Lernen [Scaffolding in inquiry learning]. *Zeitschrift für Didaktik der Naturwissenschaften, 23*(1), 21–37.

Barke, H. (2006). *Chemiedidaktik: Diagnose und Korrektur von Schülervorstellungen [Chemistry education: Diagnosis and correction of student ideas]*. Heidelberg: Springer.

Belland, B. R., Glazewski, K. D., & Richardson, J. C. (2008). A scaffolding framework to support the construction of evidence-based arguments among middle school students. *Educational Technology Research and Development, 56*(4), 401–422.

Bjonness, B., & Kolsto, S. D. (2015). Scaffolding open inquiry: How a teacher provides students with structure and space. *Nordina, 11*(3), 223–237.

Bolte, L. A. (1999). Using concept maps and interpretive essays for assessment in mathematics. *School Science and Mathematics*, *99*(1), 19–30.

Bortz, J., & Lienert, G. A. (2008). *Kurzgefasste Statistik für die klinische Forschung [Brief statistics for clinical research]* (2nd ed.). Heidelberg: Springer.

Brandstädter, K., Harms, U., & Großschedl, J. (2012). Assessing system thinking through different concept-mapping practices. *International Journal of Science Education*, *34*(14), 2147–2170.

Brewer, C. A., & Smith, D. (2009). *Vision and change in undergraduate biology education*. American Association for the Advancement of Science & National Science Foundation. Washington. Retrieved July 18, 2021, from https://www.jsg.utexas.edu/events/files/biology-vision-and-change.pdf

de Vos, W., & Verdonk, A. H. (1986). A new road to reactions: Part III. Teaching the heat effect of reactions. *Journal of Chemical Education*, *63*(11), 972.

DiBattista, D., & Kurzawa, L. (2011). Examination of the quality of multiple-choice items on classroom tests. *Canadian Journal for the Scholarship of Teaching and Learning*, *2*(2).

Djanette, B., & Fouad, C. (2014). Determination of university students' misconceptions about light using concept maps. *Procedia Social and Behavioral Sciences*, *152*, 582–589.

Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften [Research methods and evaluation in social and human sciences]*. Heidelberg: Springer.

Duit, R., & Treagust, D. F. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International Journal of Science Education*, *25*(6), 671–688.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. London: Sage.

Forghani-Arani, N., Cerna, L., & Bannon, M. (2019). *The lives of teachers in diverse classrooms, OECD Education Working Papers* (Report No. 198), OECD Publishing. Retrieved July 18, 2021, from https://doi.org/10.1787/8c26fee5-en

Franke–Braun, G., Schmidt-Weigand, F., Stäudel, L., & Wodzinski, R. (2008). Aufgaben mit gestuften Lernhilfen [Tasks with incremental scaffolds]. In R. Messner & W. Blum (Eds.), *Lernumgebungen auf dem Prüfstand* (pp. 27–42). Kassel: Kassel University Press.

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of Educational Research*, *82*(3), 300–329.

Garnett, P. J., & Treagust, D. F. (1990). Implications of research on students' understanding of electrochemistry for improving science curricula and classroom practice. *International Journal of Science Education*, *12*(2), 147–156.

Großschedl, J., & Harms, U. (2013). Assessing conceptual knowledge using similarity judgments. *Studies in Educational Evaluation*, *39*, 71-81.

Großschedl, J., Mahler, D., & Harms, U. (2018). Construction and evaluation of an instrument to measure content knowledge in biology: The CK-IBI. *Education Sciences*, *8*(3), 145.

Haltunen, K. (2003). Scaffolding performance in IR instruction: Exploring learning experiences and performance in two learning environments. *Journal of Information Science*, *29*(5), 375–390.

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. London: Routledge.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 99–107.

Kind, V. (2004). Beyond appearances: Students' misconceptions about basic chemical ideas (2nd ed.). Durham: Durham University. Retrieved July 18, 2021, from https://www.researchgate.net/publication/228799159_Beyond_Appearances_Students%27_Misconceptions_About_Basic_Chemical_Ideas

Kormi-Nouri, R., Nilsson, L.-G., & Ohta, N. (2005). The novelty effect: Support for the novelty-encoding hypothesis. *Scandinavian Journal of Psychology*, *46*, 133–43.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, *41*(4), 212–218.

Krause, U. M., Stark, R., & Mandl, H. (2004). Förderung des computerbasierten Wissenserwerbs durch kooperatives Lernen und eine Feedbackmaßnahme. [Promotion of computer–based knowledge acquisition through cooperative learning and a feedback activity]. *Zeitschrift für Pädagogische Psychologie*, *18*(2), 125–136.

Lambon-Ralph, M. A., Pobric, G., & Jefferies, E. (2009). Conceptual knowledge is underpinned by the temporal pole bilaterally: Convergent evidence from rTMS. *Cerebral Cortex*, *19*(4), 832–838.

Lou, Y., Abrami, P.C., & D'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research*, *71*(3), 449–521.

McClure, J. R., Sonak, B., & Suen, H. K. (1999). Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *Journal of Research in Science Teaching*, *36*(4), 475–492.

Mieg, H. A. (2019). Introduction: Inquiry-based learning – Initial assessment. In A. H. Mieg (Eds.), *Inquiry-based learning – Undergraduate research: The German multidisciplinary experience* (pp. 1–16). Cham: Springer Open.

Novak, J. D., & Cañas A. J. (2008). The theory underlying concept maps and how to construct them. Pensacola: IHMC. Retrieved July 18, 2021, from http://cmap.ihmc.us/publications/researchpapers/TheoryUnderlyingConceptMaps.pdf

Opitz, S. T., Blankenstein, A., & Harms, U. (2016). Student conceptions about energy in biological contexts. *Journal of Biological Education*, *51*(4), 427–440.

Pedaste, M., Mäeots, M., Leijen, Ä., & Sarapuu, S. (2012). Improving students' inquiry skills through reflection and self-regulation scaffolds. *Technology, Instruction, Cognition and Learning, 9,* 81–95.

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, *66*(2), 211–227.

Richardson, J. T. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, *6*(2), 135–147.

Rosenthal, R., & Fode, K. L. (1963). The effect of experimenter bias on the performance of the albino rat. *Behavioral Science, 8*, 183–189.

Saye, J. W., & Brush, T. (2002). Scaffolding critical reasoning about history and social issues in multimedia supported learning environments. *Educational Technology Research and Development, 50*(3), 77–96.

Schmidt-Weigand, F., Franke-Braun, G., & Hänze, M. (2008). Erhöhen gestufte Lernhilfen die Effektivität von Lösungsbeispielen? Eine Studie zur kooperativen Bearbeitung von Aufgaben in den Naturwissenschaften [The influence of different presentation modes of worked examples on learning]. *Unterrichtswissenschaft, 36*(4), 365–384.

Schmidt-Weigand, F., Hänze, M., & Wodzinski, R. (2009). Complex problem solving and worked examples. *Zeitschrift für Pädagogische Psychologie*, *23*(2), 129–138.

Smith, III, J. P., diSessa, A. A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences*, *3*(2), 115–163.

Sweller, J., van Marrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251–296.

Tan, D. K. Y., Koppi, A., & Field, D. J. (2016). First year agricultural science student perspectives in graduate attribute development through problem-based learning. *International Journal of Innovation in Science and Mathematics Education, 24*(1), 54–66.

The President's Council of Advisors on Science and Technology (2012). *Report to the President, engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Retrieved July 18, 2021, from https://files.eric.ed.gov/fulltext/ED541511.pdf

Thomas, L., Bennett, S., & Lockyer, L. (2016). Using concept maps and goal-setting to support the development of self-regulated learning in a problem-based learning curriculum. *Medical Teacher, 38*(9), 930–935.

van de Pol, J., & Elbers, E. (2013). Scaffolding student learning: A micro-analysis of teacher– student interaction. *Learning, Culture and Social Interaction, 2*(1), 32–41.

van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review, 22*(3), 271–296.

van de Pol, J., Volman, M., Oort, F., & Beishuizen, J. (2015). The effects of scaffolding in the classroom: Support contingency and student independent working time in relation to student achievement, task effort and appreciation of support. *Instructional Science*, *43*(5), 615–641.

Veríssimo Catarreira, S. M., Godinho Lopes, V., Casas García, L. M., & Luengo González, R. (2017). Evaluation of changes in cognitive structures after the learning process in mathematics. *International Journal of Innovation in Science and Mathematics Education, 25*(2), 17–33.

Wangdi, D., Precharattana, M., & Kanthang, P. (2020). A guided inquiry laboratory to enhance students' understanding of the law of mechanical energy conservation. *International Journal of Innovation in Science and Mathematics Education*, *28*(1), 29–43.

Whitley, E., & Ball, J. (2002). Statistics review 4: Sample size calculations. *Critical Care*, *6*, 335–341.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *17*(2), 89–100.

Zimmerman, B. J. (1998). Developing self-fulfilling cycles of academic regulation: An analysis of exemplary instructional models. In D. H. Schunk & B. J. Zimmerman, (Eds.). *Self-regulated learning: From teaching to self-reflective practice* (pp. 1–19). New York: The Guilford Press.

Appendix A

Misconceptions in pre- and posttest concept maps as a function of instruction condition (group)

Misconceptions	Treatment group (*n* = 59)		Control group (*n* = 46)	
	Pretest (%)	Posttest (%)	Pretest (%)	Posttest (%)
Chemical reaction consists of (is) energy	8.5	5.1	10.9	2.2
Chemical reaction needs gas burner	10.2	11.9	17.4	6.5
Gas burner needs energy	6.8	1.7	10.9	8.7
Heat generates energy	22.0	6.8	10.9	6.5
Heat needs energy	0	3.4	6.5	0
Energy generates heat	11.9	1.7	13.0	2.2
Energy consists of heat	5.1	0	0	0
Energy generates activation energy	3.4	0	6.5	0
Activation energy generates energy	1.7	0	2.2	2.2
Activation energy generates heat	3.4	1.7	0	0
Endothermic is exothermic	8.5	0	0	0
Endothermic emits heat	0	8.5	0	6.5
Exothermic absorbs heat	0	6.8	0	6.5
Water generates energy	13.6	8.5	23.9	8.7

Note. The values (in %) represent the proportion of students who use a corresponding misconception in their concept map.