Vreman-de Olde, G.C., de Jong, T., Gijlers, H.(2013). Learning by Designing Instruction in the Context of Simulation-based Inquiry Learning. *Educational Technology & Society*, *16* (4), 47-58.

Learning by Designing Instruction in the Context of Simulation-based Inquiry Learning

Cornelise Vreman-de Olde, Ton de Jong* and Hannie Gijlers
Science and Technology, University of Twente, PO BOX 217, 7500 AE Enschede, The Netherlands // g.c.vreman-deolde@utwente.nl // a.j.m.dejong@utwente.nl // a.h.gijlers@utwente.nl
*Corresponding author

(Submitted June 26, 2012; Revised February 06, 2013; Accepted February 28, 2013)

ABSTRACT

This study compares learning from designing instruction in the context of simulation-based inquiry learning with learning from expository teaching. The domain of instruction was the electricity domain of high-pass and low-pass filters. Participants were students from a technical vocational school. In the experimental condition ($N = 21$) students created assignments for an imaginary student to help this student to learn from a computer simulation. The LOOK-EXPERIMENT-DESIGN (LED) approach was developed to support students in designing these assignments. This support structure scaffolded students in orienting themselves in the simulation (LOOK), in performing experiments to gain more insight into the simulated domain (EXPERIMENT), and in designing assignments (DESIGN) about the simulated domain. Students in the control condition ($N = 28$) received traditional instruction. Students came from two different classes and were divided over the two conditions. After 3 two-hour lessons, all students completed a test measuring conceptual and procedural knowledge. Results showed that, in one class, students who learned by designing assignments performed significantly better on test items measuring conceptual knowledge than students who learned from traditional instruction. This was not replicated in the other class. No differences between the conditions were found for procedural knowledge.

Keywords

Inquiry learning, Simulations, Learning by teaching, Physics teaching

Introduction

Currently there is a general consensus that inquiry-based approaches to learning science that incorporate students' active investigation and experimentation are necessary to motivate students for science (Osborne & Dilon, 2008). Inquiry is the process in which students engage in the investigation of scientifically oriented questions, perform active experimentation, formulate explanations from evidence, evaluate their explanations in light of alternative explanations, and communicate and justify their proposed explanations (National Research Council, 2000). Beyond the motivational benefit, inquiry learning and its associated processes also have value of their own; inquiry learning generates knowledge and, if well supported, can be more effective than direct forms of instruction (Furtak, Seidel, Iverson, & Briggs, 2012). For these reasons, it is held that inquiry should be part of the science curriculum (e.g., National Research Council, 2000).

Contemporary technology-based approaches to science learning provide students with ample opportunities for inquiry. Technology-based environments offer simulations, games, data sets, and/or remote and virtual laboratories for students' inquiry-related use. Inquiry calls for non-linear and interactive content that these technology-based environments are able to provide, so that their technological affordances are directly used for pedagogical purposes (de Jong, 2006). Evidence is accumulating that technology-enhanced learning environments for inquiry provide students with genuinely effective learning opportunities, and large-scale studies show that these inquiry environments outperform more direct approaches to instruction on a variety of outcome measures (e.g.,Deslauriers & Wieman, 2011; Eysink et al., 2009; Marusić & Slisko, 2012).

However, these promising results only materialize when the inquiry process is structured and scaffolded (Alfieri, Brooks, Aldrich, & Tenenbaum, 2011). Effective scaffolds come in many forms. Examples include tools for creating hypotheses, data analysis tools, and tools for saving and monitoring experiments (see e.g., de Jong, 2005; Quintana et al., 2004; Zhang, Chen, Sun, & Reid, 2004). A growing number of computer-based inquiry environments have emerged that provide students with inquiry facilities together with an integrated supportive structure and scaffolds. Examples of such learning environments are: Inquiry Island (White et al., 2002); GenScope (Hickey, Kindfield, Horwitz, & Christie, 2003); SimQuest-based environments (de Jong et al., 1998); Co-Lab (van Joolingen, de Jong,

Lazonder, Savelsbergh, & Manlove, 2005); WISE (Linn, Davis, & Bell, 2004); STOCHASMOS (Kyza, Constantinou, & Spanoudis, 2011); and SCY (de Jong et al., 2012).
One approach that has not as yet been explored is to encourage learning from simulations by having students create scaffolds for other learners. This method is based on the idea of "learning by teaching", which assumes that in aiming to teach others tutors are encouraged to learn the domain very thoroughly themselves. From research on peer tutoring, we know that tutors gain knowledge from their teaching experience; due to their need to explain to and question the tutee, tutors engage in processes requiring reflection about or summarization of their own knowledge (Roscoe & Chi, 2007). This "learning by teaching" approach has been used in technology-enhanced learning in the "Betty's brain" software (e.g., Leelawong & Biswas, 2008). Here students learn by instructing a teachable agent, a graphical computer character equipped with artificial intelligence. Studies with Betty's brain show that teaching another is highly motivating and leads to better learning results than learning for yourself (Chase, Chin, Oppezzo, & Schwartz, 2009), and also that learning by teaching the agent leads to higher performance compared to traditional teaching methods (Biswas, Leelawong, Schwartz, Vye, & Teachable Agents Grp, 2005).

Another example is SimStudent; in SimStudent students learn by instructing a simulated student (Matsuda et al., 2010). Recent work has shown that students who teach SimStudent can achieve considerable and efficient knowledge gains, specifically if they have to reflect about their own teaching actions (Matsuda et al., 2012). However, there is also research that indicates that learning by teaching can hinder learning. Atkinson, Derry, Renkl, and Wortham (2000), for example, summarize a number of studies in which creating explanations for another learner was compared to creating explanations for oneself, with the overall finding that the students who prepared for teaching someone else scored lower on knowledge tests. These authors attribute this result to higher levels of anxiety and lower intrinsic motivation. They also indicate that having experience with tutoring beforehand may yield greater benefits from creating explanations for others. In other recent work the importance of preparing students for their tutoring role is also emphasized (Matsuda et al., 2011).

These lines of research are further explored in the current study, in which we compared learning by designing assignments for another (fictitious) student to complete in a computer simulation with learning from expository teaching. Within the present study an assignment is a question, the correct answer and an explanation of the answer. An example of a question in such assignment would be: *What happens to the output voltage of the filter if you double the frequency?* Normally, the question, the alternative(s), and the feedback for an assignment are designed by an instructional designer or teacher. In the current study the assignments were designed by students themselves, with the idea that they could learn from the design process.

In a previous study (Vreman-de Olde & de Jong, 2004) students designed assignments related to a computer simulation on electrical circuits. Two-thirds of the designed assignments were about calculations and definitions. One-third of the designed assignments were about the discoveries students made with the simulation, but these assignments were rather superficial and mainly described simple effects. To support students in designing assignments, we developed a paper-and-pencil design sheet, that prompted student to generate an idea, transform the idea into an assignment, and evaluate the assignment (by running it in the software environment). Students using this design sheet designed more assignments about the relations in the simulated domain than non-scaffolded students. In addition, scaffolded students more precisely described relations and provided more explanations than the non-scaffolded students. However, no differences between the two groups were found on a knowledge test (Vreman-de Olde & de Jong, 2006). These results, the review by Atkinson et al. (2000), and recent work on peer tutoring (e.g., Tsivitanidou, Zacharia, & Hovardas, 2011) suggest that students need (more) detailed scaffolding for their assignment designing activities.

In the present study, we compared an *experimental condition*, in which students designed assignments for a simulation on electrical circuits and were supported by a detailed design scaffold that guided the students through different steps (described more fully in the Method section), with a *control condition.* In the control condition students worked on the same learning content but followed traditional instruction in which the teacher used the blackboard for explanations and students completed calculation exercises. To assess students' learning outcomes a knowledge test with different types of test items was administered. We expected that the experimental group would perform better than the control group on conceptual test items measuring insight into the cause-effect relations of the examined domain because they would gain insight into those relations by designing assignments. Second, we expected that students in the control condition would perform better than the experimental group on procedural (calculation) items because of their greater amount of practice in performing calculations. Although these predictions seem straightforward, recent work shows that a focus on enhancing conceptual knowledge may also lead to an improvement in procedural knowledge (Kolloffel & de Jong, 2013). This result is explained by the phenomenon of bootstrapping (Schauble, 1996) or iterative knowledge development (Rittle-Johnson, Siegler, & Alibali, 2001), which refers to the idea that the acquisition of conceptual understanding and of procedural knowledge can in some cases mutually support and stimulate each other.

Method

Participants

Participants were 50 students from an intermediate level vocational engineering training program, average age 17 years. Students were from two intact classes coming from two different educational paths within technical vocational training, namely, Electronic Engineering (Class 1) and Automotive Engineering (Class 2). For their regular "practical lessons" the teachers had already split up each of the classes into two groups. In each class one of these pre-defined groups was assigned to the experimental condition, and the other to the control condition. The experimental condition consisted of 22 students, 12 students from class 1 and 10 from class 2. The control condition consisted of 28 students, 13 from class 1 and 15 from class 2. One student from the experimental condition (class 1) was absent during the test. As a result, the total experimental group consisted of 21 students. Students participated in the experiment with their own class and were instructed by their own teacher.

Materials

The domain and computer simulation learning environment

In this study, a SIMQUEST (van Joolingen & de Jong, 2003) application was used. One electrical high-pass filter and two low-pass filters were simulated. A low-pass filter is a circuit offering easy passage to low-frequency signals and difficult passage to high frequency signals, while a high-pass filter's task is just the opposite. Filters are built with two elements: A resistor (R) and a coil (L), or a resistor and a capacitor (C). In general, the theme of filters and the passage of signals is a difficult subject.
In designing the application, we used a series of four simulation interfaces for each of the three filters, presented in the same order for each filter. Complexity of the interfaces in the simulation was increased gradually, (see, e.g.,White & Frederiksen, 1990). Each series started with a simple interface presenting the elements of the filter, so that students could learn how the individual elements react to frequency changes. The second interface is shown in the left window of Figure 1. In the "Change variables" box the values of one or more variables can be changed. The output variables are visible in the "Results" box, in the "resistance diagram" and in the graph. The third interface focused on U_{out} and the current I for the whole frequency range. The fourth interface showed a graphical representation of the transfer function, plotting U_{out}/U_{in} as a function of the frequency.

The support

Students in the experimental condition were asked to design assignments so that an "imaginary" fellow student could learn from the simulation. For this task we created a support structure that guided the students through three consecutive steps: LOOK (orientation on the simulation), EXPERIMENT (experimentation with the simulation), and DESIGN (designing assignments). The rationale behind these steps, which we called LED, is that we want to focus students' attention on the relations that are important in the domain (Swaak, van Joolingen, & de Jong, 1998). After students have acquired knowledge, they are asked to make this knowledge explicit in developing a question, the correct answer and the explanation of that answer.

The support was presented in the simulation environment together with a set of paper-and-pencil worksheets (called LED-sheet hereafter). The online-support in the simulation consisted of assignments, tips, and overviews. This support was available in a window next to the simulation interface (Figure 1). The LED-sheets matched the structure of the on-line support; for example, if on-line support asked students to investigate a certain relation, instructions on the associated sheet supported students in making notes about their investigations.

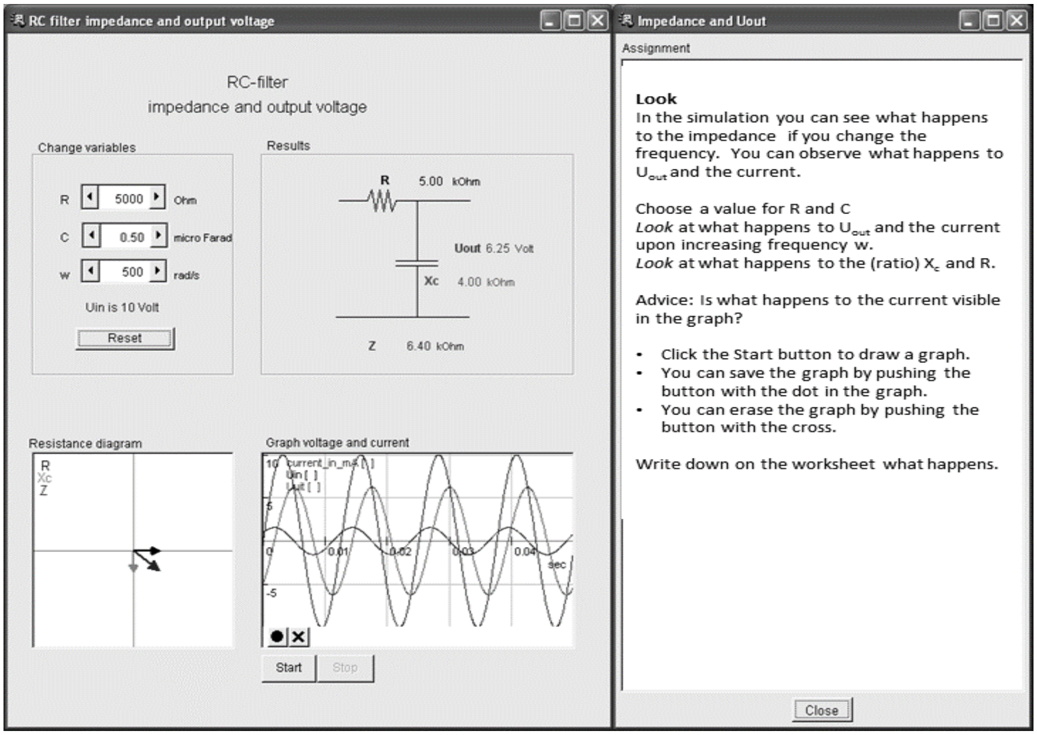

Figure 1. Screen shot of simulation interface and support from the LOOK phase.

First phase: LOOK

In the LOOK-phase, the main goal was to explore the domain. The on-line support started with an overview of the learning goals for the specific interface, followed by investigation tasks that provided students with concrete targets, so that they could perform specific inquiries in the computer environment. Students also received hints for performing experiments correctly. On the related LED-sheet, "observation starters" supported students in making notes of their observations. An observation starter is a semi-structured sentence, starting with a given focus of observation and ending with dots to be filled in. An example is: “If R increases, then…….”. By giving students this starter, observations are structured (change only R), focused (it is important to change R), and note-taking is assured (the sentence must be completed).

Second phase: EXPERIMENT

In the Experiment phase, the main goal was to transform the qualitative observations from the Look-phase into more exact descriptions. Students were supported by several detailed scaffolds.

First, students had to perform a series of systematic experiments and keep a record of the data from those measurements. To support them in this process, we included a "partly-filled-in-table" (presented in Table 1) on a LED-sheet. In this table, the (increasing) values of the independent variable and a number of the dependent variables were already given. Students had to complete this table.

Table 1. Example of the “Partly-filled-in-table”

ω (in rad/s)	X_C (in kOhm)	R (in kOhm	U_{out} (in V)	Z (in kOhm	I (in mA)
50					
100					
500					
1000					

5000					
Conclusion: If ω increases then:	Xc will ……	R will …...	U_{out} will ……	Z will …...	I will ……

The last row of the table contained what we called “conclusion starters”. These sentences were added to support students in drawing conclusions from the table. Compared with the LOOK phase the students’ statements were more precise.

Second, students were prompted to take a careful look at representations from the simulation such as formulae and diagrams. They were, e.g., asked to calculate the impedance X_C for two values of the frequency ω_1 and ω_2 (with $\omega_2 = 2\omega_1$). Students were also asked to draw diagrams (e.g., the resistance diagram in Figure 1) for different values of the frequency and draw conclusions.

Third, students were given “prediction-starters” to support them in thinking deliberately about the consequences of a change, e.g., “when the frequency becomes higher, I think the output voltage will…..”

Third phase: DESIGN

In this phase, the main goal was to design an assignment about the observations made and the knowledge acquired during previous phases. Students were supported in using this knowledge and making it explicit in their design. In generating a question, they were instructed to pose a question about the observations they had made. In formulating the answer, they were advised to check the correctness of the answer with the help of the simulation. In generating the explanation for their assignment, they were advised to explain the answer in detail, and to make use of calculations, representations, and observations. For each interface, except for the fourth one, students went through the three LED phases.

Knowledge test

Knowledge was assessed using a paper-and-pencil (post-)test. The knowledge test consisted of two parts: one set of items intended to measure conceptual (insight) knowledge, and a second set of items focused on measuring procedural (calculation) knowledge. All items were scored by a rater who was blind to the condition of the participant who had taken the test. Both the test and the answering key were developed together with the teacher.
Conceptual knowledge (insight into the cause-effect relations in the domain) was measured by items in which students were asked to predict or explain the effect of a change. Students received points for correct answers and for their reasoning. In the example shown in Figure 2, the student not only had to choose a situation, but also had to give a reason for their choice. There were a total of 28 conceptual items, with a maximum total score of 50 points; the maximum point value per item depended on its complexity (13 items with a maximum of 1 point, 9 with a maximum of 2 points, 5 with 3 points and 1 with 4 points). Reliability analysis of the test resulted in a Cronbach's alpha of 0.80. Two judges independently scored the answers to the conceptual knowledge items for ten percent of the data, with inter-rater agreement reaching 0.70 (Cohen’s kappa).

Procedural knowledge was measured by test items in which students were asked to perform calculations. Students received points for the calculation procedure and the correct answer. There were a total of 6 procedural items with a maximum total score of 15 points; the maximum point value per item depended on the its complexity (1 item question with 1 point possible, 3 with a maximum of 2 points, 2 with a maximum of 4 points). An example of a procedural item is presented in Figure 3. Reliability analysis of the test resulted in a Cronbach's alpha of 0.64. Two judges independently scored the answers to the procedural knowledge items for ten percent of the data, with inter-rater agreement reaching 0.76 (Cohen’s kappa).

There were a total of nine introductory items, that were used to “warm up” the students. These items referred to general domain knowledge and were not analyzed.

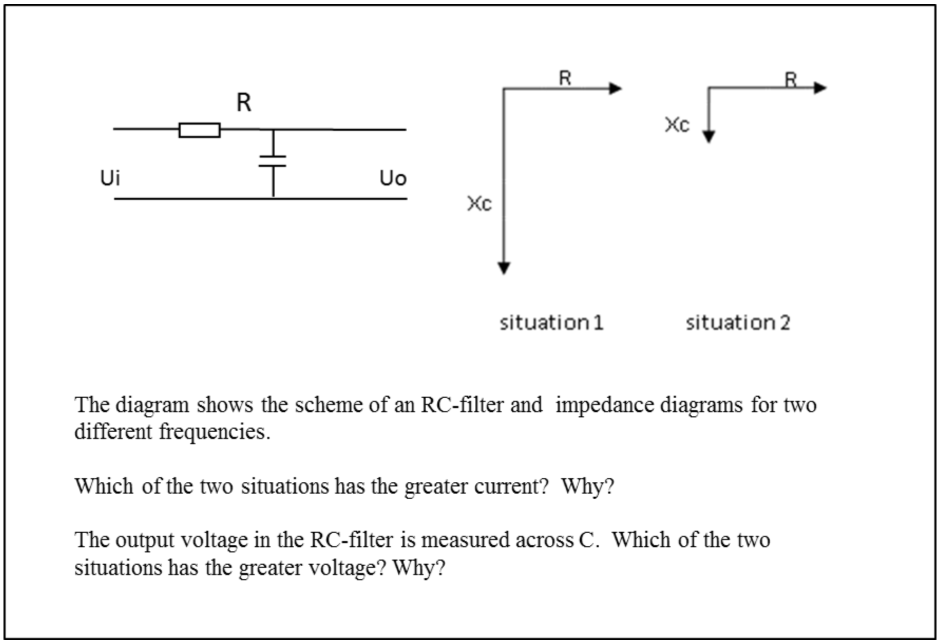

Figure 2. Example of a conceptual knowledge item

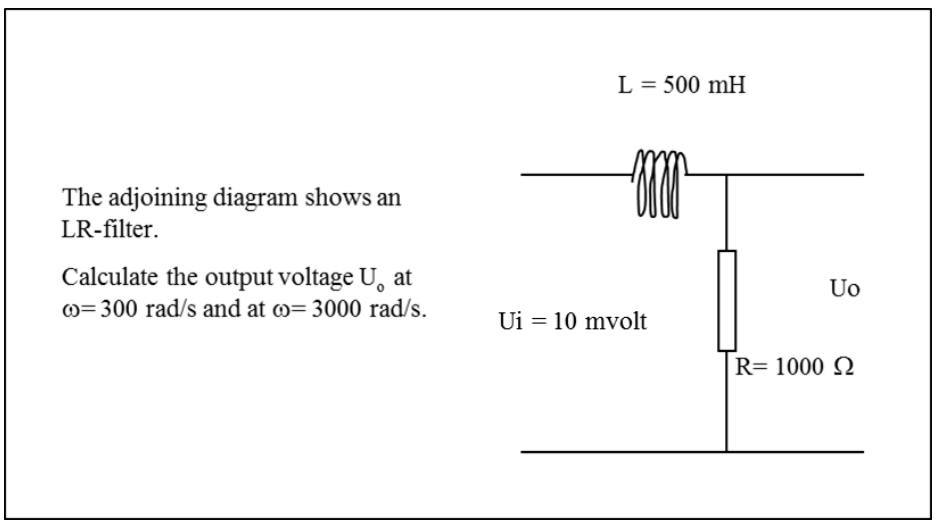

Figure 3. Example of a procedural knowledge item

Procedure

Students in both conditions had three weekly two-hour lessons on the subject of low-pass and high-pass filters, which was part of their regular curriculum. The fourth lesson was used to administer the knowledge test. Class 1 participated in the study first, and a few months later Class 2 participated. The same procedure was followed for both classes. The same domain content was covered in both conditions.

In three two-hour sessions, students in the *experimental condition* went through the simulations of each of the three filters. At the beginning of the first lesson, the experimenter introduced the students to the SIMQUEST learning environment. For the design task, the experimenter explained the three phases in the design approach and told the students how to use the LED-Sheets. During the first lesson, students worked with the simulation of the first filter. At the end of the lesson, all LED-sheets were collected. At the beginning of the second and the third lessons, the LED-sheets were returned to the students and students continued where they had stopped the lesson before. Near the end of the third lesson, students were asked to have a look at the transfer functions of each filter (they were not supposed to design assignments about transfer functions). At the end of the third lesson, LED-sheets were collected. For both classes the students' own teacher was available during all lessons to answer students' questions.

In the *control condition*, students received three two-hour lessons, from their own teacher. They did not use a computer simulation but received conventional instruction. The teacher used the blackboard for explaining the domain and students completed calculation exercises from their textbook. Informal observations of activities in the class were made during all lessons by the experimenter.

Results

In the results section, we first present the exam scores for both conditions in each class, as a way to establish the comparability of the experimental and control groups in terms of prior domain knowledge. Next, we present the results of the knowledge post-tests. Finally, to gain understanding in the way the students used the scaffolds, we analyzed students' completed paper-and-pencil design sheets.

Exam scores

Table 2 gives an overview of the mean exam scores on the subject of electricity for both conditions and for each class. The exam scores (which could range from 1-10) for this subject were made up of the scores on a number of tests from the students' regular curriculum that they had taken before the experiment began. The way these exam scores were determined in both classes was not similar (both followed a different curriculum) but data for each condition were normally distributed for both Class 1 (Experimental: Shapiro-Wilk $W = 0.959$, $df = 11$, $p = .754$; Control: Shapiro-Wilk $W= 0.969$, $df = 13$, $p = .754$) and Class 2 (Experimental: Shapiro-Wilk $W = 0.915$, $df = 10$, $p = .316$; Control: Shapiro-Wilk $W= .928$, $df = 15$, $p = .252$). No difference between the experimental and control condition was found (Class 1: $F(1,22) = .623$, $p = .438$) Class 2: $F(1,23) = .439$, $p = .514$). From this we may conclude that the experimental and control condition entered the experiment with comparable prior knowledge.

Table 2. Mean exam scores for the two classes; control and experimental condition

Condition	Experimental	Control
	M (*SD*) n (min-max)	*M* (*SD*) n (min-max)
Class 1	5.9 (*1.2*) 11 (3.7 – 7.8)	5.4 (*1.9*) 13 (2.2 – 8.4)
Class 2	6.4 (*1.6*) 10 (4.1 – 8.9)	6.8 (*1.5*) 15 (4.1 – 8.6)

The knowledge test

Table *3* shows the mean scores on the conceptual items and the procedural items from the post-test for both conditions. Shapiro-Wilk tests showed that the test scores were normally distributed for both conditions for both the conceptual items (Experimental: Shapiro-Wilk $W= .972$, $df = 21$, $p = .611$; Control: Shapiro-Wilk $W = .971$, $df = 28$, $p = .611$) and the procedural items (Experimental: Shapiro-Wilk $W= .908$, $df = 21$, $p = .051$; Control: Shapiro-Wilk $W = .947$, $df = 28$, $p = .165$). The results of a MANOVA on students' scores on conceptual and procedural items showed no significant effect of condition (F (2, 46) = 2.552, $p = .089$).

Table 3. Mean scores for the two conditions (all students) on the knowledge test

Condition	Experimental (n = 21)	Control (n = 28)
Knowledge items	*M* (*SD*)	*M* (*SD*)
Conceptual items (max = 50)	29.7 (*8.5*)	26.7 (*8.0*)
Procedural items (max = 15)	7.2 (*4.1*)	8.4 (*4.4*)

Because students came from two different classes with different backgrounds, we performed an analysis of the results on the knowledge tests for the two classes separately.

Class 1. Table 4 shows the results of the knowledge tests for the two conditions in Class 1. Statistical tests for detection of outliers showed that one student in the experimental condition appeared to be an outlier for the conceptual items (with a score greater than 2 SD below the mean score). This student was removed from further analyses. In both conditions results on the conceptual test remained normally distributed after the removal of the outlier (Experimental condition: Shapiro-Wilk W = 0.939, df = 10, p = .542; Control condition: Shapiro-Wilk W = .899, df = 13, p = .129). The test results for the procedural knowledge test were affected by the removal of the outlier (Experimental condition: Shapiro-Wilk W = .790, df = 10, p = .011; Control condition: Shapiro-Wilk W = .950, df = 13, p = .597). Therefore, the non-parametric Kruskal-Wallis test was used to examine the differences between conditions within Class 1. A significant difference between the two conditions was found on conceptual items (H(1) = 5.044, p = .025) but not on procedural items (H(1) = -.117, p = .732).

Table 4. Mean scores for the two conditions of Class 1 on the knowledge test

Condition	Experimental (n = 10)	Control (n = 13)
Knowledge items	*M* (*SD*)	*M* (*SD*)
Conceptual items (max = 50)	33.9 (*7.8*)	25.2 (*6.5*)
Procedural items (max = 15)	8.5 (*4.6*)	8.0 (*4.9*)

Class 2. Table 5 shows the results of the knowledge test for the two conditions in Class 2. Statistical tests for detection of outliers showed that one student in the experimental condition appeared to be an outlier for the conceptual items (with a score greater than 2 SD above the mean score). This student was removed from further analysis. The distribution of results remained normal for both the conceptual test (Experimental condition: Shapiro-Wilk W = 0.972, df = 9, p = .913; Control condition: Shapiro-Wilk W = .937, df = 15, p = .341) and the procedural knowledge test (Experimental condition: Shapiro-Wilk W = .970, df = 9, p = .899; Control condition: Shapiro-Wilk W = .935, df = 15, p = .318). A one-way MANOVA showed no significant differences (F(2, 21) = 2.099, p = .148) between the two conditions on the knowledge test (conceptual and procedural items). This implies that the significant difference found in Class 1 on conceptual knowledge items is not duplicated in Class 2.

Table 5. Mean scores for the two conditions of Class 2 on the knowledge test

Condition	Experimental (n = 9)	Control (n = 15)
Knowledge items	*M* (*SD*)	*M* (*SD*)
Conceptual items (max = 50)	26.1 (*4.3*)	28.0 (*9.1*)
Procedural items (max = 15)	5.7 (*2.9*)	8.8 (*4.0*)

Qualitative analysis of the support structure

The goal of our intervention was that students would create assignments with answers and explanations that would help them better understand the domain themselves. Two examples of assignments that were created by students are given in Figure 4. To gain insight into the effect of scaffolds used (see also Method section), we examined the correctness and performed a qualitative and informal analysis of the students' notes students on the LED-sheets. The results of this analysis are presented in this section.

RC-filter	
Question: Answer: Explanation:	What happens to U_{out} upon increasing R? U_{out} becomes smaller. U_{out} becomes smaller because the voltage across R becomes larger and C 'gets' less voltage.
CR filter	
Question: Answer: Explanation:	Is this a high-pass or low-pass filter? High-pass The higher the frequency, the lower the 'resistance of C', so the voltage across R becomes larger. So, this is a high-pass filter.

Figure 4. Two examples of assignments created by students

Observation starters

We hypothesized that observation starters would support students in making careful observations in the simulation and in taking notes about their observations. The answers to the observation starters were generally correct, implying that students changed the correct independent variable and observed the dependent variables we wanted them to study. In addition, we saw that some students went beyond the focus of the observation starters and described their own observations. For example, they expressed their surprise about the effect of the frequency on the impedance diagram.

Partly-filled-in-table and conclusion starters

Our assumption was that the partly-filled-in-tables would support students in performing systematic experiments, keeping an overview of their data and drawing conclusions about the relations examined. Analysis of the filled in tables shows that students filled in correct measurements of the dependent variables and that conclusions concerning the examined relations were generally correct.

Representations (1): Calculations and conclusion starter

We hypothesized that the combination of calculations on the LED sheet and conclusion starters would help students to gain a more quantitative understanding of a causal relation (see example in Method section). We found that students' answers to the calculations were mostly correct, as well as the quantitative relations they formulated.

Representations (2): Diagrams and conclusion starter

We hypothesized that drawing diagrams on the LED sheet would support students in gaining a deeper understanding of the represented concepts. Regarding the resistance diagram (see Figure 1) students were asked to draw the diagram for two values of the frequency ω. For increasing frequencies, all students drew a shorter arrow for X_C and Z (both are correct). Some students added notes about the Pythagorean formula - this formula can be used to calculate Z from X_C and R. In addition, we saw that students started to draw representations on the LED-sheets when they made their own observations. Students also drew representations when explaining their assignment answer.

Prediction starters

Lastly, we hypothesized that prediction starters would support students in thinking about effects of changes in e.g., the frequency, and in reflecting on the correctness of their predictions. In analyzing the notes made on the LED-sheets, we found that students' predictions often were not correct. We also found no reflections about the correctness of the predictions.

Discussion

In this study, we compared two version of the same learning environments. In the *experimental condition*, students designed assignments for fictitious students in a simulation based inquiry learning environment. In their design task, students went through the three design phases LOOK-EXPERIMENT-DESIGN. This sequence may not be novel to inquiry learning in general, but it is novel when used in combination with designing assignments for a computer simulation for (fictitious) other students. Along the way, students were supported such that they could gain more insight into the simulated domain. In the *control condition*, students did not use a computer simulation but received conventional instruction. The teacher used the blackboard for explaining the domain and students completed calculation exercises from their textbook. A knowledge test was administered to measure learning differences between conditions. This test contained procedural items, measuring knowledge of calculation procedures, and conceptual items, measuring insight into causal relations. The two classes that participated in our study were each divided into two groups; one group in each class participated in the experimental condition and the other group in the control condition.

Overall, we found no differences on the conceptual and procedural knowledge test items between the two conditions. Looking at the two classes separately, we found that students in the experimental condition of Class 1 performed significantly better on the conceptual items than students in the control condition. However, this result was not repeated for Class 2. This might have been caused by the difference in prior computer simulation experience between the two classes. During their regular lessons, students in Class 1, who came from a different educational program (Electronic Engineering) than students from Class 2 (Automotive Engineering), had used the program Multisim to build and simulate circuits. Experience with the Multisim simulations might have helped the students from Class 1 in learning from a simulation. For students in Class 2 this was their first encounter with computer simulations and we know from research that it takes students a bit of time to have enough experience in inquiry learning (Hickey et al., 2003; Ketelhut, 2007).

To guide the students in the experimental condition we developed a support structure. Asking students to design assignments already structures students' inquiry process. Additional scaffolds were added to support the design process so that students would be able to explore the domain, perform experiments, gather data for the design of their assignments and formulate assignments based on their newly acquired knowledge. During the whole process, students made notes on their LED-sheets. An informal qualitative analysis of these notes and the observations we made during the lessons showed that the effects of these scaffolds look promising.

First, the scaffolds "observation starter" and "drawing representations" seemed to assist students in starting an investigation, as was reflected in students' notes LED-sheets and observations made during the lessons. In addition, the representations assisted students in formulating explanations for their assignments. Drawing representations of the total impedance in the circuit helped students remember the Pythagorean formula used to calculate the total impedance. Second, the "partly-filled-in table" (developed to support students in planning and monitoring a series of experiments) assisted students in performing a series of measurements and drawing conclusions about the collected data. It seems that this relatively straightforward table helped students to keep an overview of their data and enabled them to focus on the relations being investigated. This is in line with studies that have emphasized the importance of monitoring support (Veermans, van Joolingen, & de Jong, 2006). Third, in complying with the scaffold telling them to "perform calculations", students performed two calculations with a formula and used the outcomes to describe the relation between the variables. With hindsight, we could have exploited this scaffold more by linking the calculations and its resulting quantitative relation from the EXPERIMENT phase more explicitly with the (qualitative) observations of the same relation in the LOOK phase. In this way, students might have realized that careful observations in a simulation can be used to check their understanding of a formula. A point of concern, however, is that, as our qualitative analysis showed, many predictions as formulated in the "prediction starters" were not correct. Students were not inclined to reflect on the correctness of their predictions - a process which might have given rise to interesting learning moments. Students might need extra support in reflecting on their work. This is important because reflection on one's own knowledge is a pivotal aspect of learning with computer simulations (Smetana & Bell, 2012).

The current study had limitations in the small number of participants involved and the unfamiliarity of working with simulation software for a number of the participants. This study therefore is only a first step towards designing effective "learning by design" instruction (see also de Jong et al., 2012) . The current work indicates that designing assignments for a simulation with the LOOK-EXPERIMENT-DESIGN approach opens opportunities for students to gain insight into the simulated domain. Prior experience in working with simulations seems to be a potential facilitating condition in this learning process. With respect to the scaffolds, we found that these straightforward types of support were relatively easy to use and seemed to assist our students in focusing on important relations in de domain, learning from the simulation and in using the knowledge they had acquired in designing their assignments.

References

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology, 103*, 1-18. doi: 10.1037/a0021017

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research, 70*, 181-214. doi: 10.2307/1170661

Biswas, G., Leelawong, K., Schwartz, D., Vye, N., & Teachable Agents Grp, V. (2005). Learning by teaching: A new agent paradigm for educational software. *Applied Artificial Intelligence, 19*, 363-392. doi: 10.1080/08839510590910200

Chase, C., Chin, D., Oppezzo, M., & Schwartz, D. (2009). Teachable agents and the protégé effect: Increasing the effort towards learning. *Journal of Science Education and Technology, 18*, 334-352. doi: 10.1007/s10956-009-9180-4

de Jong, T. (2005). The guided discovery principle in multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 215-229). Cambridge, UK: Cambridge University Press.

de Jong, T. (2006). Computer simulations - technological advances in inquiry learning. *Science, 312*, 532-533. doi: 10.1126/science.1127750

de Jong, T., van Joolingen, W. R., Swaak, J., Veermans, K., Limbach, R., King, S., & Gureghian, D. (1998). Self-directed learning in simulation-based discovery environments. *Journal of Computer Assisted Learning, 14*, 235-246. doi: 10.1046/j.1365-2729.1998.143060.x

de Jong, T., Weinberger, A., Girault, I., Kluge, A., Lazonder, A. W., Pedaste, M., . . . Zacharia, Z. C. (2012). Using scenarios to design complex technology-enhanced learning environments. *Educational Technology Research & Development, 60*, 883-901. doi: 10.1007/s11423-012-9258-1

Deslauriers, L., & Wieman, C. E. (2011). Learning and retention of quantum concepts with different teaching methods. *Physical Review Special Topics - Physics Education Research, 7,* 010101. doi: 10.1103/PhysRevSTPER.7.010101

Eysink, T. H. S., de Jong, T., Berthold, K., Kollöffel, B., Opfermann, M., & Wouters, P. (2009). Learner performance in multimedia learning arrangements: An analysis across instructional approaches. *American Educational Research Journal, 46*, 1107-1149. doi: 10.3102/0002831209340235

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching. *Review of Educational Research, 82*, 300-329. doi: 10.3102/0034654312457206

Hickey, D. T., Kindfield, A. C. H., Horwitz, P., & Christie, M. A. (2003). Integrating curriculum, instruction, assessment, and evaluation in a technology-supported genetics environment. *American Educational Research Journal, 40*, 495-538. doi: 10.3102/00028312040002495

Ketelhut, D. (2007). The impact of student self-efficacy on scientific inquiry skills: An exploratory investigation in River City, a multi-user virtual environment. *Journal of Science Education and Technology, 16*, 99-111. doi: 10.1007/s10956-006-9038-y

Kolloffel, B., & de Jong, T. (2013). Conceptual understanding of electrical circuits in secondary vocational engineering education: Combining traditional instruction with inquiry learning in a virtual lab. *Journal of Engineering Education, 102*, 375-393.

Kyza, E. A., Constantinou, C. P., & Spanoudis, G. (2011). Sixth graders' co-construction of explanations of a disturbance in an ecosystem: Exploring relationships between grouping, reflective scaffolding, and evidence-based explanations. *International Journal of Science Education, 33*(18), 1-37. doi: 10.1080/09500693.2010.550951

Leelawong, K., & Biswas, G. (2008). Designing learning by teaching agents: The Betty's brain system. *International Journal of Artificial Intelligence in Education, 18*, 181-208.

Linn, M. C., Davis, E. A., & Bell, P. (2004). Inquiry and technology. In M. Linn, E. A. Davis & P. Bell (Eds.), *Internet environments for science education* (pp. 3-28). Mahwah, NJ: Lawrence Erlbaum Associates.

Marusić, M., & Slisko, J. (2012). Influence of three different methods of teaching physics on the gain in students' development of reasoning. *International Journal of Science Education, 34*, 301–326. doi: 10.1080/09500693.2011.582522

Matsuda, N., Cohen, W. W., Koedinger, K. R., Keiser, V., Raizada, R., & Yarzebinski, E. (2012). Studying the effect of tutor learning using a teachable agent that asks the student tutor for explanations. In M. Sugimoto, V. Aleven, Y. S. Chee & B. F. Manjon (Eds.), *International conference on digital game and intelligent toy enhanced learning (digitel 2012)* (pp. 25-32). Takamatsu (Japan): Los Alamitos, CA: IEEE Computer Society.

Matsuda, N., Keiser, V., Raizada, R., Tu, A., Stylianides, G., Cohen, W., & Koedinger, K. (2010). Learning by teaching SimStudent: Technical accomplishments and an initial use with students. In V. Aleven, J. Kay & J. Mostow (Eds.), *Intelligent tutoring systems* (Vol. 6094, pp. 317-326). Berlin: Springer.

Matsuda, N., Yarzebinski, E., Keiser, V., Raizada, R., Stylianides, G., & Cohen, W. W. (2011, January). *Learning by teaching SimStudent - an initial classroom baseline study comparing with cognitive tutor*. Paper presented at the Proceedings of the International Conference on Artificial Intelligence in Education.

National Research Council. (2000). *Inquiry and the national science education standards. A guide for teaching and learning.* Washington, DC: National Academy Press.

Osborne, J., & Dilon, J. (2008). *Science education in Europe: Critical reflections*. London, United Kingdom: Nuffield Foundation.

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *The Journal of the Learning Sciences, 13*, 337-387. doi: 10.1207/s15327809jls1303_4

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93*, 346-362.

Roscoe, R. D., & Chi, M. T. H. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. *Review of Educational Research, 77*, 534-574. doi: 10.3102/0034654307309920

Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology, 32*, 102-119. doi: 10.1037/0012-1649.32.1.102

Smetana, L. K., & Bell, R. L. (2012). Computer simulations to support science instruction and learning: A critical review of the literature. *International Journal of Science Education, 34*, 1337-1370. doi: 10.1080/09500693.2011.605182

Swaak, J., van Joolingen, W. R., & de Jong, T. (1998). Supporting simulation-based learning; the effects of model progression and assignments on definitional and intuitive knowledge. *Learning and Instruction, 8*, 235-252. doi: doi:10.1016/S0959-4752(98)00018-8

Tsivitanidou, O. E., Zacharia, Z. C., & Hovardas, T. (2011). Investigating secondary school students' unmediated peer assessment skills. *Learning and Instruction, 21*, 506-519. doi: 10.1016/j.learninstruc.2010.08.002

van Joolingen, W. R., & de Jong, T. (2003). Simquest: Authoring educational simulations. In T. Murray, S. Blessing & S. Ainsworth (Eds.), *Authoring tools for advanced technology educational software: Toward cost-effective production of adaptive, interactive, and intelligent educational software* (pp. 1-31). Dordrecht, The Netherlands: Kluwer Academic Publishers.

van Joolingen, W. R., de Jong, T., Lazonder, A. W., Savelsbergh, E., & Manlove, S. (2005). Co-lab: Research and development of an on-line learning environment for collaborative scientific discovery learning. *Computers in Human Behavior, 21*, 671-688. doi: 10.1016/j.chb.2004.10.039

Veermans, K. H., van Joolingen, W. R., & de Jong, T. (2006). Using heuristics to facilitate scientific discovery learning in a simulation learning environment in a physics domain. *International Journal of Science Education, 28*, 341-361. doi: 10.1080/09500690500277615

Vreman-de Olde, C., & de Jong, T. (2004). Student-generated assignments about electrical circuits in a computer simulation. *International Journal of Science Education 26*, 859-873. doi: 10.1080/0950069032000138815

Vreman-de Olde, C., & de Jong, T. (2006). Scaffolding the design of assignments for a computer simulation. *Journal of Computer Assisted Learning, 22*, 63-74. doi: 10.1111/j.1365-2729.2006.00160.x

White, B. Y., Frederiksen, J., Frederiksen, T., Eslinger, E., Loper, S., & Collins, A. (2002, October). *Inquiry island: Affordances of a multi-agent environment for scientific inquiry and reflective learning.* Paper presented at the Fifth International Conference of the Learning Sciences (ICLS), Seattle, WA.

White, B. Y., & Frederiksen, J. R. (1990). Causal model progressions as a foundation for intelligent learning environments. *Artificial Intelligence, 42*, 99-157.

Zhang, J., Chen, Q., Sun, Y., & Reid, D. J. (2004). Triple scheme of learning support design for scientific discovery learning based on computer simulation: Experimental research. *Journal of Computer Assisted Learning, 20*, 269-282. doi: 10.1111/j.1365-2729.2004.00062.x