Investigating a Method of Scaffolding Student-Designed Experiments

Kelly Morgan · David W. Brooks

Abstract The process of designing an experiment is a difficult one. Students often struggle to perform such tasks as the design process places a large cognitive load on students. Scaffolding is the process of providing support for a student to allow them to complete tasks they would otherwise not have been able to complete. This study sought to investigate backwards-design, one form of scaffolding the experimental design process for students. Students were guided through the design process in a backwards manner (designing the results section first and working backwards through typical report components to the materials and safety sections). The use of reflective prompts as possible scaffold for metacognitive processes was also studied. Scaffolding was in the form of a computer application built specifically for this purpose. Four versions of the computer application were randomly assigned to 102 high school chemistry students and students were asked to the design of an experiment, producing a report. The use of backwards-design scaffolding resulted in significantly higher performance on lab reports. The addition of reflective prompts reduced the effect of backwards-design scaffolding in lower-level students.

Keywords Scaffolding · Inquiry · Assessment

K. Morgan (✉)
Kansas State University, Manhattan, KS, USA
e-mail: kellymorganscience@gmail.com

D. W. Brooks
University of Nebraska-Lincoln, Lincoln, NE, USA
e-mail: dbrooks@unlserve.unl.edu

Introduction and Literature Review

Documents and science standards have been urging instructors to use inquiry-based education, including activities in which students design laboratory experiments, for over 15 years (AAAS 1993, 2001; NRC 1996). A survey of 571 high school chemistry teachers found only 55.5% of the respondents utilized student-designed experiments in their courses. One possible explanation for the lack of student-designed experiments in the classroom is the inability of students to undertake such tasks successfully in a reasonable amount of time. When presented with a task as difficult as designing their own experiment, students often do not know where to begin or where to go next (Morgan Deters 2006). There is a large continuum of inquiry ranging from guided to open inquiry. However, teachers often feel that, in order to be regarded as "inquiry learning," the students must be guided minimally, if at all. Minimally guided, or open inquiry, instruction can be frustrating to both student and teacher and is often unsuccessful (Kirschner et al. 2006).

Critics of inquiry approaches generally adhere to cognitive load theory wherein three types of cognitive load are placed on working memory: intrinsic load, extraneous load and germane load (Van Merrienboar et al. 2006). Intrinsic load refers to the task difficulty. Extraneous load results from inefficient/inappropriate instructional design and is not necessary for the task to be completed. Germane load involves creating or altering schema or chunks of information. The forms of cognitive load are additive. If the intrinsic and extraneous load of a task meets or exceeds the capacity of the learner's working memory, there will be no room for germane load and meaningful learning is unlikely. Kirschner et al. (2006) demonstrate that student-designed investigations have high cognitive load and that the

literature providing results of greater learning gains through minimally guided instruction is actually providing case studies in which very effective teachers seamlessly scaffold students when they fail to make progress in the open discovery setting. The positive effects in such case studies cannot be attributed to minimal guidance (open inquiry), but instead to scaffolded or guided inquiry.

As the presence of scaffolding or guidance is necessary for students to successfully undertake experimental design processes, a reliable, effective method of scaffolding is sought. For this study, scaffolding took place in the form of a computer application created specifically for the purpose of guiding students through the process of designing an experiment (Morgan Deters 2009).

Scaffolding

Although students may have experience reading and performing an investigation that is provided to them, they may have few or no current schema relevant to the design process. The result is an overload on working memory resources during design tasks. To overcome this large burden on processing ability, scaffolding is needed for students in novel problem-solving tasks (Kirschner et al. 2006; Shell et al. 2010). Scaffolding is "the precise help that enables a learner to achieve a specific goal that would not be possible without some kind of support" (Sharpe 2006). Many authors have connected the term "scaffolding" with "zone of proximal development" (for example, Bruner 1985; Holton and Clarke 2006; McNeill et al. 2006; Sharpe 2006; Shepard 2005). Vygotsky described the zone of proximal development in 1930 as "the distance between the actual developmental level as determined by independent problem-solving under adult guidance or in collaboration with more capable peers. The zone represents the potential for a child's development when aided by others" (Vygotsky 1978, p. 86).

Scaffolding is the assistance a learner needs, carefully balanced between too much support resulting in the task becoming to easy and too little support resulting in cognitively overloaded and frustrated learners, to succeed in the zone of proximal development. As the zone of proximal development shifts with the learner acquiring knew schema or skills, the scaffolding must also shift, change or be reduced to accommodate the new zone of proximal development and the new tasks.

Computerized Scaffolding

One-on-one scaffolding is not always a viable option in a classroom with a single instructor (for example: Davis 2003; McNeill et al. 2006; Puntambekar and Kolodner 2005). Therefore this study sought to investigate a computerized scaffolding tool for student-designed experiments.

Quintana et al. (2004) envisioned software itself as the scaffolding rather than using "help functions" and other options within software as the scaffolding. Sharpe (2006) discussed reducing the degrees of freedom of carrying out a task, allowing students to concentrate on acquiring knowledge while reducing the cognitive load of unproductive paths. Quintana et al. (2005) gave four specific recommendations for scaffolding of this type: provide a work space for what's required, provide progress displays (on computerized scaffolding), display information important to a process each that that information is important and display the "big picture focal point" (such as the research question) on each screen, page or workspace.

Backwards Design

A pre-designed experiment typically has characteristic components: Problem/Purpose, Materials and Safety, Procedure, Data, Results, Analysis, and sometimes Discussion of Errors. Consider the cognitive load implied by creating a design in this sequence. In order to determine what is appropriate in the Materials and Safety section of the investigation report, students would need to know what they are going to do in the investigation. In order for them to know what they are going to do, they must know what data or observation they will need to collect. Before they know what data or observations are important, students need to know what calculations, comparisons or trend recognitions will be needed to appropriately address the problem or purpose of the investigation. The early step in a conventional approach to experiments, creating a list of materials and safety concerns, is in reality available only at the end of a multi-step process in which experiments are designed.

However it is possible that designing an experiment in a different order may be more efficient and effective. The first task is to identify information that sheds light on the question at hand. In other words, what data do we need? Then we decide which measurements or observations might yield that data. Next we ask how to make those observations or measurements. The world of experimentation—in terms of designing experiments—runs in reverse from the sequence often used when reporting those experiments.

For several years we have tried to design classroom inquiry scaffolding that reflects this more realistic design sequence. It was our hypothesis that a "backwards-design" method requires less information to be held in working memory at each step in the design process. For example, when designing the procedure for the experiment the student would be able to reference the calculations/results

set-up as well as the data table they have already written down rather than holding that information in their working memory, as would be necessary if they were designing the procedure before the other sections. From a cognitive load perspective, therefore, it should be easier to accomplish. This led us to conduct experiments in which the "backwards design" was compared with the conventional design. The hypothesis was that the scaffolding, specifically in the form of backwards design, would lead to improved learning when compared with conventional design.

Reflective Prompts

If the intention of scaffolding is help students to develop skills and schema to allow them to complete such tasks in the future, then the scaffolding must at some point begin to fade. In order for students to develop the necessary schema related to allow the fading of the scaffolding, they must pay specific attention to the design process itself (Shell et al. 2010). Students must be aware of what steps they are taking and why. This type of metacognitive process can have two goals: (1) to determine if learners have allocated appropriate working memory to the task and (2) to encourage the storing of the new knowledge or skills in a manner in which they may be more easily called upon in future tasks (Brooks and Shell 2006). Students need to be prompted while in the "flow" or "zone" to reflect in addition to including reflection time at the completion of a task (Brooks and Shell 2006; Norman 1991). Puntambekar and Kolodner (2005) implemented a "Design Diary" to study methods of scaffolding self-regulatory skills and demonstrated this need to interrupt students when "in the flow of work" by being prompted to reflect. If students need to reflect throughout and following the task, what is the most effective way to prompt the reflection?

Specific and contextualized prompts are often used to promote reflection, however a generic prompt, one which is not specific to when in the process it is provided, may be more effective for a wide variety of student ability. Davis (2003) found students that were classified as poor reflectors performed significantly better when provided with generic reflective prompts whereas students that were better reflectors showed no difference between generic and specific reflective prompts. She suggested that the specific prompts may not be within the zone of proximal development of the poor reflecting students and therefore were not successful supports for them. Winne (1995) and Manlove et al. (2007) have also presented evidence that the type of support for reflection or self-regulatory practices may increase the cognitive load of the learner and therefore decrease the quality of either the primary task or the responses to the reflective prompts.

The scaffolding used in this study sought to determine the effect of generic reflective prompts on student performance while designing an experiment with the hypothesis that generic reflective prompts, as opposed to non-use of reflective prompts, would increase student performance on the task.

Research Questions

1. How does the backwards-design scaffolding affect the quality of the student investigation reports?
2. How do reflective prompts affect the quality of student investigation reports?

Methods

Setting

The setting of the study was a Midwestern, public, suburban/rural high school with an enrollment of 1,112 freshman through seniors. The sample for this study consisted of sophomores, juniors and seniors in a general chemistry course.

In this school district, eighth grade science teachers place students in freshman science courses. Top students enroll in biology as freshmen. Mid-level students are placed in a pre-chemistry lab-intensive course, while low-average to low-level students enroll in a semester of physical science and a semester of earth/space science. Top students generally proceed to chemistry during the sophomore year; other students tend to take the course as juniors.

Participants

Participants were 102 students in high school general chemistry courses. Table 1 displays student demographics.

A quasi-experimental 2 × 2 design was used to study the two variables: effect of backwards-design process and effect of reflective prompting. The treatment design is show in Table 2.

Of the 138 general chemistry students, 118 returned signed consent and assent forms, 103 of whom were present the day of the activity. Participants were randomly

Table 1 Participant demographics, number of students

Gender	Year in school	Freshman year course
Male—48	Sophomores—7	Physical science/earth science—58
Female—54	Juniors—82	Biology—13
	Seniors—13	Empirical science—30

Table 2 Treatment design

Backwards-designed process	Student-determined design process
Reflective prompts	Reflective prompts
n = 29	n = 28
Backwards-designed process	Student-determined design process
No reflective prompts	No reflective prompts
n = 22	n = 23

assigned a username and password as well as a random assignment of treatment group. After participants were provided with login information, the spreadsheet connecting student names with login information was deleted. The username, password and software version assigned were stored in a MySQL database with the experiment reports generated throughout the inquiry task. Classrooms remained in tact and students accessed the software through individual desktops in one of the school's computer labs.

Participant Task

The participant task was to design an investigation to determine which of two brands of paper towel absorbed the most water per dollar cost. This investigation appears in the Kendall Hunt Chemistry: Discovering Chemistry You Need to Know high school textbook (Morgan Deters 2008). As a result of classroom use, modifications have been made to increase clarity of directions and instructions over the years. Although the science content of this task is low and students have extensive experience with paper towels in everyday life, this task presents a challenge to students each year. Students in this school have had very little, if any, experience designing their own experiments in previous science courses. Therefore the designing of the investigation is enough to produce cognitive overload for most students in this sample, as evidenced by poor performance and the need for a great amount of assistance and direction during this task in previous years.

Each student was presented with the same investigation instructions and grading rubric (see Online Resource) through the computer software. Assistance from the teacher was limited to help navigating the software (for example: which button to click next, how to save their information, etc.).

Instruments, Measures and Analysis

An investigation report rubric (see Online Resource) was developed by examining the author's current rubric for scientific investigations, the National Science Education Standards (NRC 1996), Benchmarks for Science Literacy

(AAAS 1993) and the Atlas of Science Literacy (AAAS 2001) for appropriate performance levels on investigation design for high school students. Skills and proficiencies sited in these documents were used to determine the categories and criteria for the grading rubric. The rubric has been validated by high school chemistry teachers in multiple locations giving comments and suggestions for modification following use of the rubric with their own students. Rubric score reliability was determined with a random sample of 20% of the reports being graded by another chemistry teacher in the school.

Each lab report was assessed according to this rubric. ANOVA calculations were used to determine main effects of backwards-design scaffolding, of reflective prompt presence and interaction effects between the two variables on the student lab report scores as determined by the rubric.

Scaffolding Software

Students were presented with one of four versions of the Student-Designed Labs Scaffolding and Assessment Tool (SDL-SAT) created for this study (Morgan Deters 2009). The original SDL-SAT was a web-based interface created with php (a programming language for dynamic web pages, http://www.php.net) and storing all student input and task-specific instructions from the teacher in a MySQL database (an open-source relational database, http://www.mysql.com). A new user-interface was programmed using Runtime Revolution (a programming tool for creating interactive interfaces, http://www.runrev.com), allowing much greater flexibility in design and programming. During the school year of this experiment, eight teachers used the interface for a total of 750 student experiment reports for 24 different assigned tasks.

The student interface allows students to log into his/her account and select an activity that has been assigned to them by their teacher. Teachers create and edit the activities by customizing instructions for each component of the task and including their own rubric grading information. Each screen the students view shows only the information important to that step in the design process (to lower extraneous information that may cause cognitive overload Clark and Mayer 2003).

For example, when developing the materials list the students do not need to view the data table and results sections of the lab report, simply the procedure section. Presenting all teacher instruction and rubric information for only the current student field also eliminated extraneous information. For example, when a student moved the cursor from the results set-up field to the data table design field, the screen would clear of any results section instructions and rubric information and display only the instructions and rubric information for only the data table

section of the report. Although students could move back and forth throughout the steps using on-screen buttons, the steps are arranged in the backwards-design manner within the SDL-SAT.

For the purpose of this study, an experimental version of the SDL-SAT was created. Students logged into the system using randomly assigned usernames and passwords, and were randomly assigned to one of the four versions. The four versions of the SDL-SAT were:

- Backwards-design process with reflective prompts (BDP-R)

    - The layout and presentation of the minimum necessary information was as described above (See Fig. 1). Students were guided through the process in a one screen at a time to scaffold the order in which components were designed. The screens were:

        - Screen 1 (shown in Fig. 1): title, purpose, background information and results section set-up
        - Screen 2: data table design (students could still view purpose and results sections on this screen)
        - Screen 3: procedure development (students could still view purpose and data table sections on this screen)
        - Screen 4: materials and safety (students could still view purpose and procedure sections on this screen).

    - After completing screens 1, 2 and 3, students were presented with a textbox and a prompt to "reflect on

the progress you have made towards the goal of the lab up to this point." These responses were also stored in the database with the student lab reports.

- Backwards-design process without reflective prompts (BDP)

    - This version was the same as BDP except students were not prompted to reflect.

- Student-determined process with reflective prompts (SDP-R)

    - The student-determined process versions allowed the students to determine the order in which they design report sections (See Fig. 2).
    - This version most closely resembled pencil and paper work as students were presented with all of the sections of the lab report on one screen and were not guided in any way (such as the BDP version which took one step at a time in a specific order) as to which section to complete next. Paper and pencil was not chosen as a treatment in this study as students could not have the same experience of having instructions and rubric information appear for only the section they're currently working on with paper and pencil. Also, students would have had more than one piece of paper to look at (the paper they were using and the paper with the instructions, rubric information). It was thought the this treatment lead to higher cognitive load due to the Split-Attention Effect (Clark et al. 2006).
    - After completing all lab components on the single screen, students were presented with a textbox and a

Fig. 1 Screenshot of SDL-SAT. As students click in a textbox on the left half of the screen to enter in their lab report information, the instructions and grading rubric are displayed for that lab report section in the text box on the right side of the screen

Fig. 2 Screenshot of SDP version of the software. Rather than being led through the process one screen at a time to scaffold the order in which components were designed, students were presented with all lab components on one screen with a scrollbar and allowed to design in any order
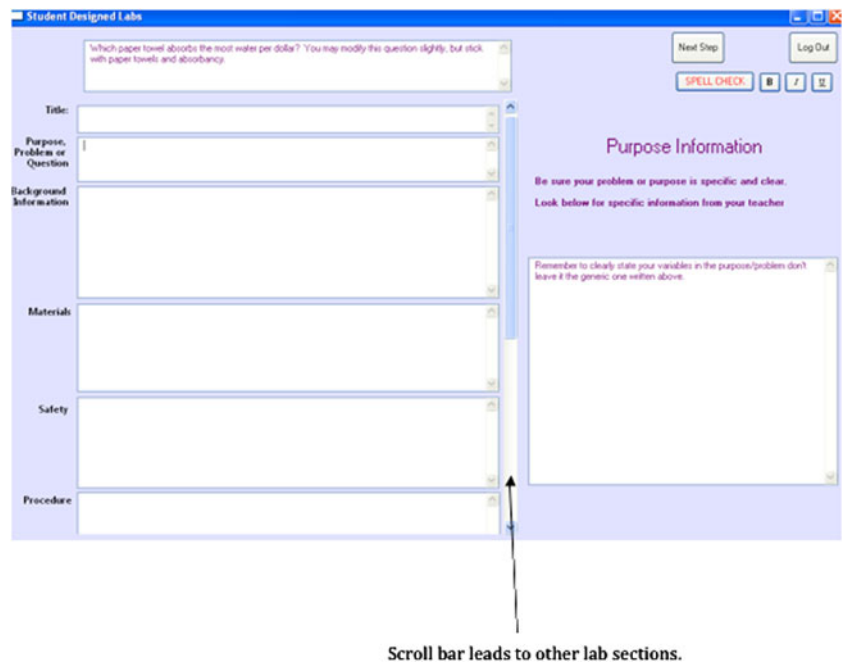
Scroll bar leads to other lab sections.

prompt to "reflect on the progress you have made towards the goal of the lab up to this point." These responses were also stored in the database with the student lab reports.

- Student-determined process without reflective prompts

  - This version was the same as SDP-R without the reflective prompts.

Procedure

The activity took place within the first 2 weeks of the school year. The timing of the task and the use of the same curriculum across different classes minimized curricular and teacher differences in this study.

On the first day of school, all students were presented with an IRB consent and assent letters. Students and parents were informed that all students were required to complete this lab task and it was a normal part of the curriculum. Students that did not return the consent and assent forms completed the task on paper. All students received a participation grade for the task to ensure that no student grades were adversely affected by which version of the SDL-SAT software they were randomly assigned.

Students typically work in pairs to complete these types of tasks in the classroom, however we felt allowing students to work in pairs during this study would add many uncontrolled factors, such as the dynamic nature of the pairing and the quality of collaboration between the pair. Therefore it was decided to have students work individually on this task.

Students were instructed on the overall task goal of finding the most absorbent paper towel per dollar between two brands. The classes remained in the computer lab until all students were ready to gather their data (within the first class period of 100 min). The second 100-min class period was spent in the classroom/laboratory performing their experiments, gathering data, analyzing data and writing conclusions. Students finished gathering data at different times and were able to rotate through the six classroom computers to finish their lab reports within the SDL-SAT software.

Each student's activity was assessed using the teacher interface of the SDL-SAT (See Fig. 3). This displayed a student's lab report side-by-side with the rubric scoring information. Each section of a lab report was scored according to the rubric (see Supplementary Material) and the scores for each section were stored in the database along with the student lab report.

Results

Of the 102 participants, only 83 of them completed the task within the designated two class periods due to absences and extra-curricular activities. All 102 students did complete the planning phase of the task within the first class period. The complete lab report score, as determined by the grading rubric, was correlated with the score for the only the sections completed in the first class period for those that completed the task in the designated two class periods. For these 83 students, the planning score (first class period)

Fig. 3 Screenshot of teacher interface. Student lab report is shown in the textboxes on the left. When the teacher clicked on a specific lab report section, the instructions and rubric that had been shown to the students was displayed on the right. Each section had a list of possible scores that was selected by the teacher. The computer application automatically totaled the scores for each lab and stored them in the database along with the scores for each individual section
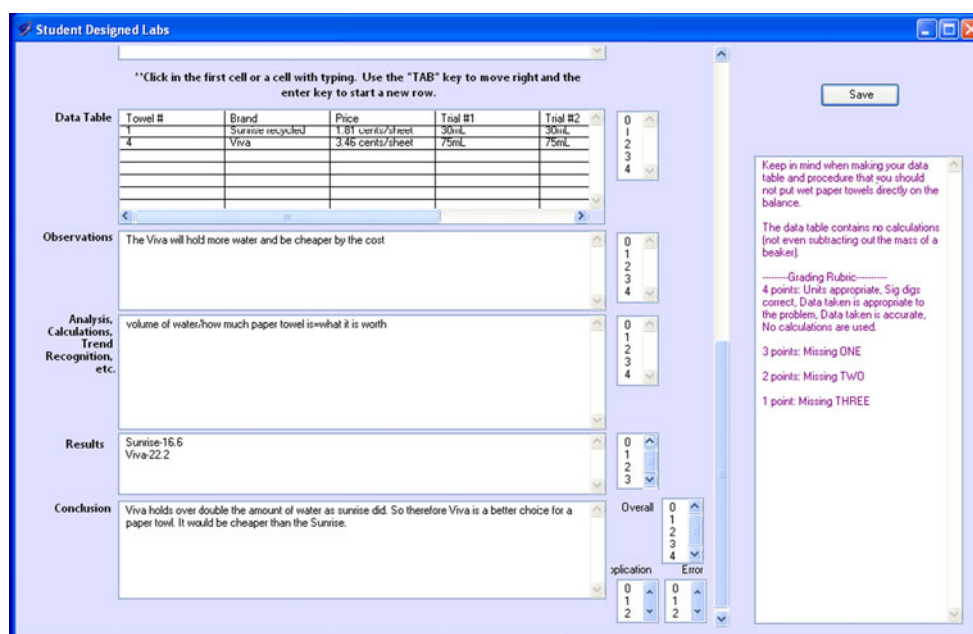
correlated with their total grade (over both class periods) with a Pearson correlation of r = 0.872 ($p = 0.00$). With such a strong correlation between the planning phase which all students completed and the entire lab report, all 102 students' planning scores were used to maximize the sample size.

A random sample of 20 reports was chosen from the pool of 102 reports. The other chemistry teacher in the school then scored the selected reports. The Pearson correlation between the planning score awarded by the author and the other chemistry teachers was r = 0.872 ($p = 0.00$), a strong inter-rater correlation suggesting consistent use of the rubric for scoring.

Bartlett's test of variance resulted in a non-significant finding, $p = 0.997$, ensuring equal variances in all treatment groups.

Table 3 summarizes the scores for all 102 participants on the tasks completed in the first class period.

A significant interact effect was found with ANOVA, $F(1,98) = 4.127$, $p = 0.045$. Simple main effects were then studied.

When no reflective prompts were used, the difference in score for student-determined design order (M = 19.09) and student-determined design order (M = 22.45) was significant, $F(1,43) = 7.26$, $p = 0.010$, with a Cohen's d effect size of 0.82, a large effect (Cohen 1977). When students were not prompted to reflect, students scaffolded with the backwards-design method significantly out-performed students free to choose their own order of design on lab report scores.

However, when students were prompted to reflect, there was virtually no difference in backwards-design

Table 3 Summary of student planning scores

Student planning scores	No reflective prompts	Reflective prompts
Backwards design scaffolding	M = 22.45	M = 20.48
	s = 4.056	s = 4.128
	n = 22	n = 29
Student-determined design order	M = 19.09	M = 20.50
	s = 4.133	s = 4.350
	n = 23	n = 28

Maximum score possible: 30 points

scaffolding versus student determined design order with means of 20.48 and 20.50 respectively. The significant effect of design order scaffolding disappeared when students were prompted to reflect on their progress with an effect size of only 0.0045, no effect.

Discussion

Backwards-Design Scaffolding

When presented with an experimental design challenge, even with low science content as in this task, a high cognitive load is placed on students of this level, as evidenced by an average grade of 68.7% on this task without aide from the teacher or peers. This is consistent with finding discussed by Kirschner et al. (2006). We hypothesized that backwards-scaffolding would have a positive effect on student lab report scores.

Kirschner et al. discussed one type of scaffolding, known as "process worksheets," that guides students through steps of a task by asking questions or prompting students for the next step. Process worksheets have been shown to be an effective way to scaffold students and increase student performance on cognitively loaded tasks. The backwards-design scaffolding in this study is similar to process worksheets in that it guides students to the next step (e.g. after designing the data table, the next step is to write the procedure). The backwards-design scaffolding gave a similar result in student performance to the process worksheets, as seen in the two treatment groups without reflective prompts where backwards-design students outperformed unguided students with a large effect size of 0.82.

Reflective Prompts

Process worksheets or backwards-design prompts can scaffold an experimental design task. Similarly, self-regulatory tasks can also be scaffolded. Students can be prompted to reflect on their progress and many papers discuss the need for reflective scaffolding (Puntambekar and Kolodner 2005; Brooks and Shell 2006; Norman 1991, etc.). Davis (2003) found generic prompts are more effective for students classified as poor reflectors and there is no difference in effectiveness for directed versus generic prompts with high-level reflectors. We hypothesized that the use of generic prompts throughout the design process would increase student performance on the written lab report.

Table 3 shows the use of generic reflective prompts in two of the four versions of the software in an attempt to scaffold reflection. The presence of reflective prompts reduced the dramatic effect of backwards-design scaffolding ($p = 0.01$, $d = 0.82$) to zero effect ($p = 0.988$, $d = 0.0045$). Students in the backwards-design scaffolding were prompted to reflect three times as they progressed through the various screens during the experiment design process. However, the student-determined design order students were only prompted to reflect one time after they had designed all of the planning sections of the lab report that were all present on one screen to allow them to design in any order. Manlove et al. (2007) suggested that too much interaction with supports or scaffolds can negatively affect mental model formation. Winne (1995) also presented arguments of increased cognitive load for regulatory processes. Both Manlove et al. and Winne presented arguments that this effect may be more of a cause for concern for lower-level students.

To determine if the multiple reflective prompts caused increased cognitive load for low-level students in this study, the performance scores must be disaggregated by student level. The only way to differentiate students by

Table 4 Summary of student planning scores for advanced students only

Plan scores for advanced students	No reflective prompts	Reflective prompts
Backwards design scaffolding	M = 24.18	M = 23.71
	s = 2.04	s = 3.039
	n = 11	n = 7
Student-determined design order	M = 19.23	M = 20.92
	s = 4.456	s = 4.582
	n = 13	n = 12

Maximum score possible: 30 points

overall academic level in this study was by the science course taken in their freshman year. Higher-level students would have been placed in either biology or the pre-chemistry course their freshman year. Table 4 displays student performance scores for higher-level students in this study.

For scores of advanced students only, the interaction effect between reflective prompts and design order scaffolding disappears, $F(1,39) = 0.811$, $p = 0.373$. For these students, there is a main effect for design order scaffolding, $F(1,39) = 10.494$, $p = 0.002$, which is a very large effect ($d = 1.5$). Reflective prompts had no effect for advanced students, $F(1,39) = 0.259$, $p = 0.613$. This disappearance of the interaction effect and large main effect of design order scaffolding for both reflective prompted and non-prompted students when considering advanced student performance scores support the suggestion that prompting lower-level students to reflect multiple times may interfere with other task processes.

Table 5 compares the findings when analyzing all student scores versus advanced student scores.

Manlove et al. (2007) demonstrated that the presence of help systems negatively affected low-level students. Winne (1995) found that generic reflective prompts were more beneficial to low-level students. This study showed that increasing the number of times students are presented with generic prompts negatively affects student performance for low-level students. These three studies provide evidence that the presence, type and frequency of reflective prompts and other help systems affect higher-level and lower-level students differently. Winne (1995) suggest that these forms of self-regulatory scaffolding be implemented after students have begun to develop schema and prior knowledge relevant to the main task to free working memory space for self-regulatory scaffolding.

Limitations of the Study

Several limitations of this study should be kept in mind. First, students knew they were receiving participation

Table 5 Summary of effects for all students and advanced students

	Interaction effect	Backwards scaffolding	Reflective prompts
All students	Significant effect	Significant effect, positive large effect size ($p = 0.010$, $d = 0.82$) for no-reflective prompts only	No significant effect
Advanced students	No significant effect	Significant positive large effect ($p = 0.002$, d = 1.5)	No significant effect

grade only on this task and therefore were not motivated to complete their best work.

Although students were placed at individual computers in a computer lab with constant teacher supervision, it was difficult to maintain the individual nature of the task once students entered the lab to gather their data. Students worked at lab tables with other students and often discussed what they were doing with each other. The effect of access to peer interactions during this task can be seen by comparing the previous year's scores on this activity to those in this study. During the previous school year, the average score on this task was 84.55% with a standard deviation of 8.15%. During that year, students worked in pairs or a group of three and had access to individualized help from the teacher. For the experiment described in this paper, the average score was 68.7% with a standard deviation of 14.2%. This effect was lessened by using the planning scores rather than scores that included sections of the lab report completed after they had the opportunity to interact with each other concerning this task.

A larger sample size would have provided larger disaggregated sample sizes of advanced and lower-level students. A large number of advanced students at the school enroll in Advanced Chemistry and it was decided that using students in a different course would introduce too many other variables. This decision resulted in a smaller sample of advanced students for the study.

Students can become more proficient at the software through multiple exposures and therefore further studies should repeat this process for multiple experimental design tasks in a course.

Conclusion

Unguided experimental design tasks prove to be high cognitively loaded for students, yet students designing their own investigations are incorporated in science education standards. Effective methods to scaffold schema development for these skills and tasks are needed to make student-designed experiments viable and effective in classrooms. This study has shown backwards-design scaffolding to increase student performance scores and the effect of reflective prompts to be student ability level dependent.

In its current form, this software does not provide the type of individualized, faded scaffolding a skilled teacher can provide. However, this study does serve as the foundation for a scaffolding and assessment software that can be studied in the future to increase its ability to support students, increase germane cognitive load and track student progress towards inquiry goals set in place by standards. This form of computerized scaffolding has shown promise in aiding students as they design investigations. Future studies can begin to create a dynamic system that adjusts the level of scaffolding, both for the design process and self-regulatory functions such as reflection, as appropriate for the student based on past student performance and the difficulty of the task being undertaken.

References

American Association for the Advancement of Science (AAAS) (1993) Benchmarks for science literacy. Oxford, New York

American Association for the Advancement of Science (AAAS) (2001) Atlas for science literacy. American Association for the Advancement of Science, Washington

Brooks DW, Shell DF (2006) Working memory, motivation and teacher-initiated learning. J Sci Educ Technol 15(1):17–30

Bruner J (1985) Vygotsky: a historical and conceptual perspective. In: Wertsch JV (ed) Culture, communication & cognition. Cambridge University Press, Cambridge

Clark RC, Mayer R (2003) e-Learning and the science of instruction. Pfeiffer, San Francisco

Clark R, Nguyen F, Sweller J (2006) Efficiency in learning: evidence-based guidelines to manage cognitive load. Pfeiffer, San Francisco

Cohen J (1977) Stastical power analysis for behavioral sciences. Academic Press, San Diego

Davis E (2003) Prompting middle school science students for productive reflection: generic and directed prompts. J Learn Sci 12(1):91–142

Holton D, Clarke D (2006) Scaffolding and metacognition. Int J Math Educ Sci Technol 37(2):127–143

Kirschner PA, Sweller J, Clark R (2006) Why minimal guidance during instruction does not work: an analysis of the failure of

constructivist, discovery, problem-based, experiential and inquiry-based teaching. Educ Psychol 41:75–86

Manlove S, Lazonder AW, de Jong T (2007) Software scaffolds to promote regulation during scientific inquiry to promote regulation during scientific inquiry learning. Metacogn Learn 2:141–155

McNeill K, Lizotte DJ, Krajcik J (2006) Supporting students' construction of scientific explanations by fading scaffolding in instructional materials. J Learn Sci 15(2):153–191

Morgan Deters K (2006) What are we teaching in high school chemistry? J Chem Educ 83(10):1492–1498

Morgan Deters K (2008) Kendall/hunt chemistry: discovering chemistry you need to know. Kendall/Hunt, Dubuque

Morgan Deters, K. (2009). Student-designed laboratory scaffolding and assessment tool (Version 3.0.1.0). Retrieved from www.studentdesignedlabs.com

National Research Council (1996) National science education standards. National Academy Press, Washington

Norman DA (1991) Cognitive artifacts. In: Carroll JM (ed) Designing interaction. Cambridge University Press, Cambridge, pp 17–38

Puntambekar S, Kolodner J (2005) Toward implementing distributed scaffolding: helping students learn science from design. J Res Sci Teach 42(2):185–217

Quintana C, Reiser B, Davis E, Krajcik J, Fretz E, Duncan R (2004) A framework for supporting metacognitive aspects of online inquiry through software-based scaffolding. Educ Psychol 13: 337–386

Quintana C, Zhang M, Krajcik J (2005) A framework for supporting metacognitive aspects of online inquiry through software-based scaffolding. Educ Psychol 40(4):235–244

Sharpe T (2006) 'Unpacking' scaffolding: identifying discourse and multimodel strategies that support learning. Lang Educ 20(3): 211–231

Shell DF, Brooks DW, Trainin G, Wilson KM, Kauffman DF, Herr LM (2010) The unified learning model: how motivational, cognitive, and neurobiological sciences inform best teaching practices. Springer, New York

Shepard L (2005) Linking formative assessment to scaffolding. Educ Leaders 63(3):66–70

Van Merrienboar J, Kester L, Paas F (2006) Teaching complex rather than simple tasks: balancing intrinsic and germane load to enhance transfer of learning. Appl Cogn Psychol 20(3):343–352

Vygotsky LS (1978) Mind in society: the development of higher psychological processes. In: Cole M, Scriber S, Johns-Steiner V, Souberman E (eds) Harvard University Press, Cambridge (original work published 1930)

Winne PH (1995) Inherent details in self-regulated learning. Educ Psychol 30(4):173–184