J Sci Educ Technol (2017) 26:33–43
DOI 10.1007/s10956-016-9649-x

The Effect of Scaffolding Strategies for Inscriptions and Argumentation in a Science Cyberlearning Environment

Cindy L. Kern[1] · Kent J. Crippen[2]

Published online: 30 August 2016

Abstract Scientific inscriptions—graphs, diagrams, and data—and argumentation are integral to learning and communicating science and are common elements in cyberlearning environments—those involving the use of networked learning technologies. However, previous research has indicated that learners struggle to use inscriptions and when they engage in argumentation, the learning of science content becomes secondary to the learning of argumentation skills. The purpose of this study was to evaluate two scaffolding strategies for these elements in a secondary school context: (1) *self-explanation prompts* paired with a scientific inscription and (2) *faded worked examples* for the evaluation and development of scientific arguments. Participants consisted of ninth and tenth grade students (age 13–16 years; $N = 245$) enrolled in state-mandated biology courses taught by four different teachers. A three-factor mixed model analysis of variance with two between factors (self-explanation prompts and faded worked examples) and one within factor (pre-, post-, delayed posttest) was used to evaluate the effects on the acquisition and retention of domain-specific content knowledge. Results indicated that neither strategy influenced the acquisition and retention of science content in a positive (i.e., learning) or negative (i.e., expertise reversal effect) way. Thus, general prompts were as effective as either of the scaffolding conditions. These unanticipated results suggest that additional research is warranted for learning scaffolds with pre-college populations where the gains were established with college-aged participants.

Keywords Cyberlearning · Design · Learning environments

✉ Cindy L. Kern
Cindy.Kern@quinnipiac.edu

Kent J. Crippen
kcrippen@coe.ufl.edu

[1] School of Education, Quinnipiac University, North Have, CT, USA

[2] School of Teaching and Learning, University of Florida, Gainesville, FL, USA

Introduction

Scientists and science students express their thinking through a variety of written and oral forms. Though scientists communicate with content rich inscriptions—graphs, diagrams, and data—the interpretation of these representations by students, those who do not have an equivalent level of understanding, is problematic (Glazer 2011). Scientists also engage in the practice of collaborative evidence-based discourse called argumentation, which is the keystone to the generation of scientific knowledge (Bricker and Bell 2008). Over the past 20 years, researchers have asserted that students who engage in argumentation are learning how scientific knowledge is generated while also learning science content. However, the existing research supporting this claim has focused almost entirely on strategies for enhancing the quality of learner's arguments with little regard to the effect of argumentation on a learner's acquisition of content knowledge (Cross et al. 2008; Ruiz-Primo et al. 2002; von Aufschnaiter et al. 2008).

Scientific inscriptions and argumentation are common elements in cyberlearning environments—those involving the use of networked learning technologies (Martinez and Peters Burton 2013). One of the advantages attributed to learning in this way is the capacity to provide technology-enhanced supports (i.e., scaffolding) that allows learners to engage in tasks that are beyond their independent abilities (Kim and Hannafin 2011), such as those involving the use of inscriptions and argumentation. Though empirical research supports the general strategy of scaffolding (e.g., van Merriënboer et al. 2003), in situ research involving specific forms of scaffolding in a technology-enhanced secondary science classroom setting is lacking.

The 5-Featured Dynamic Inquiry Enterprise (5-DIE) is a research-based framework that supports the design of inquiry-based cyberlearning in secondary science classrooms (Kern and Crippen 2013). 5-DIE provides comprehensive guidelines and includes a range of supports and strategies that are deployed in order to promote critical thinking, problem solving, and content area learning through inquiry (Kern et al. 2014). The purpose of this study was to evaluate the effect of two scaffolding strategies in a 5-DIE lesson—one for inscriptions and the other for argumentation—on the acquisition and retention of content knowledge for ninth and tenth grade students enrolled in state-mandated biology course. The following research questions guided our study:

1. What effect does learning with inscriptions paired with self-explanation prompts in a cyberlearning lesson have on the acquisition and retention of content knowledge?
2. What effect does learning with faded worked examples in a cyberlearning lesson have on the acquisition and retention of content knowledge?
3. What combined effect do these learning strategies have on the acquisition and retention of content knowledge?

The development of 5-DIE is based upon design-based research (DBR), and this study represents the third iteration of our DBR process (McKenney and Reeves 2012). Next, we describe cognitive load theory as the theoretical rationale for scaffolded learning. This is followed by a review of research supporting our use of self-explanation prompts and faded worked examples as scaffolding strategies for cyberlearning.

Review of Related Research

Cognitive load theory (CLT) has emerged over the past 20 years as an influential explanation for learning and subsequent rationale for the design of instruction. The primary claim of this theory is that without considering a learner's cognitive architecture, the effectiveness of instructional design is likely to be random (Schnotz and Kürschner 2007). In terms of a learner's ability to store and process information, CLT assumes that working memory is restricted to approximately seven chunks of information (Sweller 1988) and is constrained further when processing new or complex information (van Gog et al. 2008). Theoretically, long-term memory can store an unlimited amount of organized knowledge (Kalyuga et al. 2003). Therefore, the goal of learning strategies is to help the learner develop specific schema in long-term memory that are organized, categorized, and automated for a problem solution.

The term cognitive load refers to any demand on working memory and the processing of information (Schnotz and Kürschner 2007), and there are three types experienced by learners: (1) intrinsic, (2) extraneous, and (3) germane. Intrinsic cognitive load refers to the interplay that occurs between long-term and working memory, or the influence prior knowledge has on working memory—the greater the prior knowledge, the lower the intrinsic cognitive load. Cognitive load is also affected by external factors such as the influence of instructional design on working memory. Extraneous cognitive load refers to the demand instructional design places on working memory with processing that is unrelated to schema development or the organization of knowledge about a specific concept (van Merriënboer et al. 2003). Germane cognitive load occurs when instructional design engages learners in processing that leads to the development of a cognitive schema or influences the storing of information in long-term memory (Kalyuga et al. 2003; Sweller 1988). Germane and intrinsic cognitive load are more desirable than extraneous load.

Cognitive load theory helps us understand why presenting novices with new concepts and procedures for problem solving results in extraneous cognitive load and, therefore, little learning because the abundance of new information taxes working memory (Kalyuga 2009). Carefully designed lessons that support the engagement of a learner by reducing extraneous cognitive load and fostering germane cognitive load result in increased learning, increased learning efficiency, and greater depth of understanding (Atkinson et al. 2000). Therefore, selection of an appropriate strategy or technique is viewed as an essential component of lesson design.

Scaffolding refers to the instructional support provided to learners that allows them to engage in a task in a productive manner beyond their independent abilities (Kim and Hannafin 2011), and any positive impact of scaffolding on learning would be attributed to shifting the memory demands from extraneous cognitive load to germane cognitive load (Schnotz and Kürschner 2007). However, as a learner's skills improve, the scaffolds should be slowly removed or faded (Atkinson et al. 2003). In the context of the current study, scaffolding is the specific affordance of the instructional design that provides cognitive and social supports developed to strengthen learner content knowledge acquisition. Specifically, we focus on conceptual scaffolds; elements that are intended to help learners to bridge the gap between what they already know and what they need to know. These scaffolds were implemented as: (a) self-explanation prompts paired with scientific inscriptions and (b) faded worked examples of the evaluation and development of a scientific argument. In the next sections, a description and illustration of the effect or advantage of each scaffolding strategy is presented, as well as descriptions of key research studies consistent with a wide body of relevant research.

Scaffolding Strategy #1: Self-Explanation Prompts

Self-explanation is a form of self-talk where a learner engages in an iterative personal dialog while problem solving (Chi et al. 1994). Self-explanation prompts are conceptual scaffolds designed to guide students through the self-explanation process as they work to understand and integrate the concept, procedure, or representation (Berthold et al. 2009). The guidance provided by self-explanation prompts reduces the extraneous cognitive load associated with unguided activities, thereby providing a direction for the cognitive processing necessary for learning (Kirschner et al. 2006).

In a review of research on the self-explanation effect, Roy and Chi (2005) describe large learning effects for self-explanation in multimedia environments. Learning outcomes increase within active learning contexts and with high levels of scaffolding. In general, when self-explanation prompts are compared to little or no scaffolding, the effect size is medium to large (Berthold et al. 2009). When specific self-explanation prompts are compared to less specific self-explanation prompts, there is a small effect (van der Meij and de Jong 2011). Therefore, it was hypothesized that learning with scientific inscriptions by prompting self-explanation during a 5-DIE lesson would result in the acquisition and retention of content knowledge. The scaffolding of a scientific inscription with self-explanation prompts should (a) shift cognitive load from extraneous to germane load and (b) make explicit the science concepts represented in the inscription, thus resulting in the acquisition and retention of content knowledge.

Scaffolding Strategy #2: Faded Worked Examples

Worked examples provide detailed solutions along with the processes used to solve a problem (Crippen and Earl 2007). The provided solution gives structure for understanding how the problem is solved without a script or procedure for the solution. Instructional design incorporating worked examples reduces cognitive load by minimizing extraneous demands, allowing the learner to focus on understanding the application of the principles in the presented solutions (Renkl et al. 2004). Extensive research has been conducted on the use of worked examples with well-structured problem solving tasks, such as those common in STEM (e.g., Sweller and Cooper 1985; van Gog et al. 2008; Kissane et al. 2008).

For learners with low prior knowledge, worked example-based instruction has shown to be effective in many STEM contexts. However, due to a process known as the expertise reversal effect, for individuals with high prior knowledge, this form of learning is less effective (Clark et al. 2012; Kalyuga et al. 2003). The expertise reversal effect occurs when added information increases extraneous cognitive load because the learner already has an existing schema for solving the problem, and working memory is required to process the redundant information.

Fading is the term used to describe the process of progressively removing solution steps in a worked example while having a learner provide the missing information (Atkinson et al. 2003). In numerous studies, this strategy has been shown to mitigate the expertise reversal effect and produce medium to large learning effects (Renkl et al. 2004; Kissane et al. 2008; Schwonke et al. 2009). Thus, faded worked examples can be applied as a conceptual scaffold to shift cognitive load from extraneous to germane load in order to improve learning. Therefore, it was hypothesized that using faded worked examples to scaffold the evaluation and development of a scientific knowledge claim during a 5-DIE lesson would result in the acquisition and retention of domain-specific content knowledge.

Methodology

Using a quasi-experimental design, participants were assigned to one of four conditions: a lesson that included neither of the scaffolding strategies (control condition), a lesson that included a scientific inscription paired with reflective self-explanation prompts (self-explanation condition), a lesson that included faded worked examples for the evaluation and development of scientific arguments (faded worked example condition), and a lesson that included both scaffolding strategies (combined condition). This design was used to assess the effect of self-explanation prompts, the effect of faded worked examples, and the effect of the combination of both scaffolding strategies on the acquisition and retention of content knowledge.

Setting

The school context for this study was a large suburban school in the southwestern USA with a 41 % ethnic minority population. Seventeen percent of the school's population qualified for free/reduced lunch, and 8 % had documented disabilities. Participants were enrolled in four Biology Honors, eleven General Biology, and three Inclusionary Biology classes taught by four different teachers. Biology Honors is a yearlong course that presents biological concepts in a rigorous manner to academically oriented students. General Biology and Inclusionary Biology are both yearlong courses. Each course is designed as a survey of the biological sciences. The general classes consist primarily of general education students (i.e., students without a documented disability). The inclusionary classes consist of students with documented disabilities (approximately 50 % of the students) and general education students. Students with learning disabilities, emotional disorders, or health issues have Individual Education Plans (IEP), and most are enrolled in the Inclusionary Biology course.

The Learning Environment and Interventions

The lesson involved what Staker and Horn (2012) define as a *flex model* of blended learning—where the Internet provides the content and instruction while the teacher serves in an adaptable role, providing as-needed support on an individual basis. The lesson was constructed using SoftChalk, a software authoring tool, and delivered using the Moodle Learning Management System (LMS). The participants worked in groups of two assigned by their classroom teachers, and each group shared one laptop computer. As the participants worked collaboratively through the lesson, they completed a formal digital product individually called a Research Brief. The Research Brief was a document template with sentence starter prompts where participants recorded their thinking, analysis, reflection, and synthesis.

During the lesson, participants explored the effect that human activity or climate change could have on organisms in the same and neighboring ecosystems. Each of the five scenarios in the lesson was based on the manipulation of an interactive computational systems model with participants collecting the output as data (Fig. 1). Models were built using STELLA software from ISEE Systems and delivered online with Forio Simulate (http://forio.com/simulate/). Regardless of the condition, participants were instructed to complete all components of each lesson.

Scaffolding Condition #1: Self-Explanation of Inscriptions

A scientific inscription paired with self-explanation prompts was provided during data collection in two of the experimental condition (with self-explanation prompts and the combined condition). The self-explanation prompts were used to elicit reflection and explanation. In response to the scenarios, using the provided scientific inscription, the participants recorded the effect a change in one population could have on other populations in the simplified food web (Fig. 2).

Scaffolding Condition #2: Faded Worked Examples for Argument Development

The development of the scientific knowledge claim for the lesson was scaffolded with a series of worked examples in which components of the arguments were removed, or faded, leaving the participants to complete the remaining tasks. Each claim statement was based on the scenarios, where the participants are asked to develop a claim statement about the effect a change in one population could have on other populations in the simplified food web (Fig. 3).

Methods

A three-factor mixed model analysis of variance (ANOVA) with two between factors and one within factor was used to allow for multiple independent variables to be systematically evaluated (Table 1). The first independent variable was a scientific inscriptions paired with reflective self-explanation prompts (with/without); the second independent variable was a faded worked example strategy for the evaluation and development of scientific arguments (with/without); and the within factor was time (pretest to posttest to delayed posttest). The participants' scores on the content knowledge instrument that is detailed below represented the dependent variable.

Participants

The participants consisted of ninth and tenth grade students (age 13–16 years; $N = 245$) enrolled in Biology Honors, General Biology, or Inclusionary Biology courses. In order to reduce the potential for a teacher effect on the outcomes, the interventions were assigned so each teacher taught each intervention. The anticipated and consented/assented number of participants was 367. However, due to attrition and a school emergency that limited several classes from participating in the posttest, the actual number of participants that were used for analysis was 245. Using a strategy of whole classroom assignment, the initial participants were assigned to one of the four conditions based on teacher and biology course: control condition ($N = 40$), self-explanation condition ($N = 44$), faded worked examples condition ($N = 61$), and combined condition with both self-explanation and faded worked examples ($N = 100$).

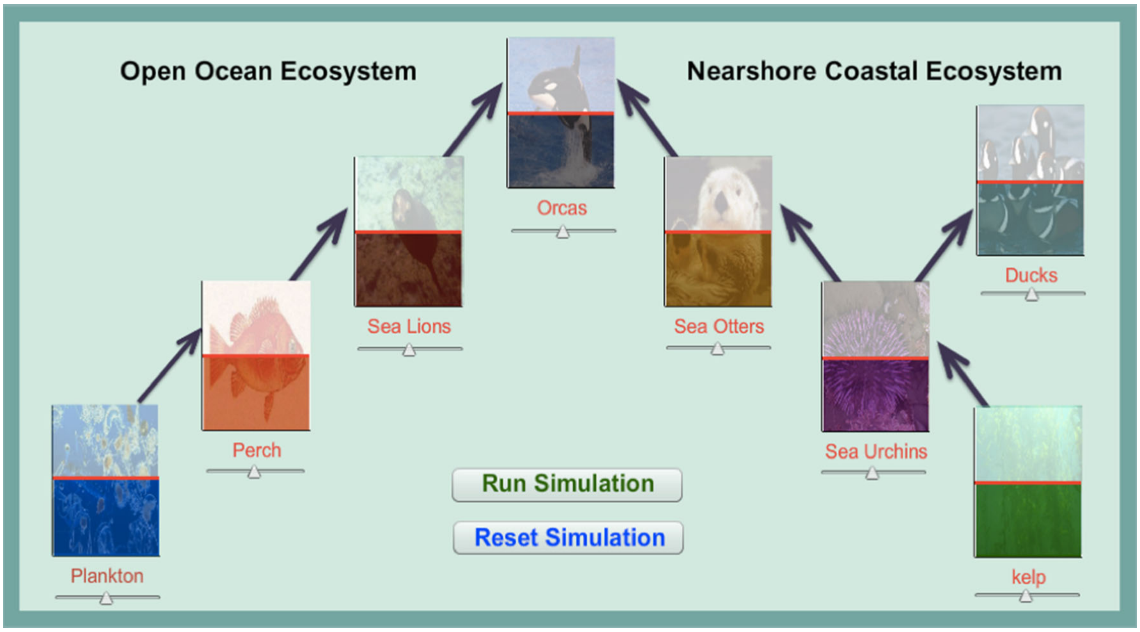
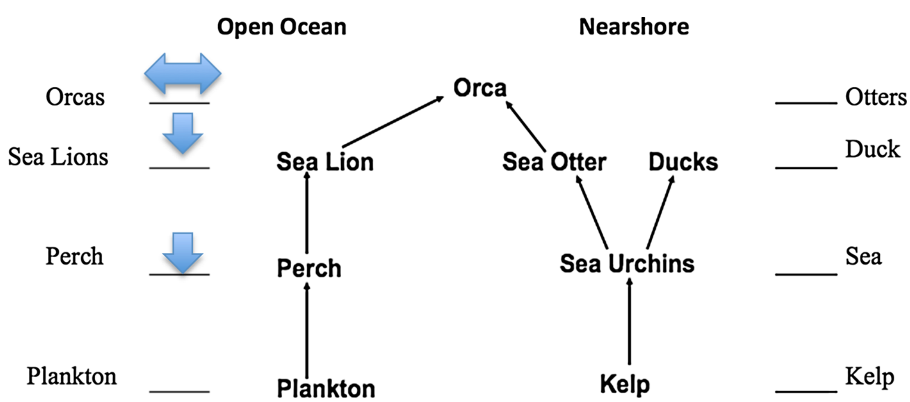


Fig. 1 Screenshot of the user interface for one interactive computational model that participants manipulated to evaluate changes in an open ocean and nearshore ecosystem. Moving the slider for one species changes the corresponding quantity of other related species (as represented by the *red level line* in each species *box*) (see http://bit.ly/2bhU8Qv) (Color figure online)

Fig. 2 Screenshot of the inscription paired with self-explanation prompts from scenario #1 of the lesson. This figure illustrates the scientific inscription paired with self-explanation prompts intervention for scenario #1 in the self-explanation prompts and combine conditions



Review the evidence you have collected and then complete the following starter prompts:

1. Two reasons the graphic organizer above is helpful in data collection are…
2. The graphic organizer helped me…
3. In the graphic organizer there are ***two*** different types of arrows that represent ***two*** different science concepts. Explain to your partner what each arrow represents then describe them below.
   a. Bold (/) arrows represent…
   b. Thin arrows (Plankton → Perch) represent…
   c. The overfishing of the perch impacted the plankton because…
   d. Two ways the nearshore ecosystem (right side of the food web) was impacted by overfishing include…

Fig. 3 Series of three faded worked examples. Worked example #1 includes a description of the individual components of a claim statement (i.e., evidence, claim, and relationship), and sample claim statement is provided. Worked example #2 requires participants to identify each to component of the argument. Worked example #3 the claim statement by stating the evidence

Worked example #1

Your Scientific Claims
A scientific claim statement has three components or parts: Evidence, Claim, and Relationship. Observations resulting in data collection are the basis of evidence. Evidence is a statement of what you *saw* happen in the model. When data is analyzed it may provide evidence to support a conclusion or **claim** for the investigation. In a scientific claim statement the relationship between the claim and evidence must be explained.

The following is an example of a scientific claim statement. Each component of the scientific claim statement is highlighted differently in order to help you see the structure.

- Evidence
- **Claim**
- Relationship

When the number of perch was decreased in our simulation the sea lion and sea otter populations decreased and there was an increase in the plankton, sea urchin, duck, and kelp populations, therefore **I claim that overfishing of perch in the open ocean impacts the energy available in the food web** because when humans remove the perch for humans to eat they remove the energy available to other animals in the food web causing changes in the ecosystems.

Worked example #2

The following is a scientific claim statement about human impact on a food web. Read the claim statement then highlight each component of a claim statement with the assigned color.

- Evidence (Blue)
- Claim (Green)
- Relationship (Red)

Human activity impacts population in a food web when they build homes and hotels where ducks nest and feed because there is an increase in the number of ducks eating sea urchins in the nearshore environment causing a change in available energy for other populations. We saw this in the model when the duck population was increased the sea urchin, sea otter, plankton and sea lion populations' decrease and the kelp and perch populations increased.

Worked example #3

The following is a scientific claim statement about human impact on energy transfer in ecosystems. Read the claim and reason parts of the statement then complete the statement by adding the *evidence that supports the claim and relationship.*

I claim that an increase in global temperatures impacts energy transfer in an ecosystem because kelp live in cold waters and warm temperatures lead to a decrease in the number of kelp reducing the amount of available energy in the food web. This is evidenced by…

Data Source

Content Knowledge Instrument

Content knowledge related to the concept of energy transfer in an ecosystem was assessed with a researcher-developed survey instrument. Through a detailed evaluation of the lesson, a table of specification was developed to identify the intended science content and the approximate percentage of the intervention dedicated to each strand (Notar et al. 2004). The content knowledge instrument was developed using items from the school district's assessment item pool, as well as modified questions from the following sources: the Program for International Student Assessment (PISA), the National Assessment of Educational Progress (NAEP), and the American Association for the Advancement of Science (AAAS) assessments. In addition to the items selected from these sources, items were developed by the researcher using a two-tiered format that required a response to a question and a statement of the reasoning for the response (Lee and Liu 2010). These assessment items were chosen because they were designed to measure content knowledge of high school students of equivalent age.

Each assessment item was categorized by the researcher and an expert panel (Rew et al. 2003) using Webb's (2007) depth of knowledge (DOK) in which the test items were considered for both the content assessed and the depth to which the learner is expected to demonstrate understanding

Table 1 Characteristics of the content knowledge instrument

Item	Depth of knowledge (DOK)	Content	Origin
1	1	Efficiency of energy transfer	District assessment
2	1	Efficiency of energy transfer	District assessment
3	2	Efficiency of energy transfer	Modified from a practice PISA
4	3	Efficiency of energy transfer	Researcher developed[a]
5	1	Evaluate the impact of changes	AAAS science assessment
6	2	Evaluate the impact of changes	District assessment
7	2	Evaluate the impact of changes	AAAS science assessment
8	3	Evaluate the impact of changes	Researcher developed[a]
9	3	Evaluate the impact of changes	District level assessment
10	1	Human impact	Researcher developed[a]
11	2	Human impact	Researcher developed[a]
12	3	Human impact	Modified NAEP practice
13	1	Transfer of matter and energy	District assessment
14	1	Transfer of matter and energy	Modified NAEP practice
15	2	Transfer of matter and energy	District assessment
16	3	Transfer of matter and energy	Researcher developed[a]
17	3	Transfer of matter and energy	AAAS science assessment

[a] Based upon the knowledge integration two-tiered assessment format (Lee and Liu 2010)

of that content (Table 1). A school district DOK specialist was consulted for the categorization. To mitigate the potential for a ceiling effect, each content strand had one assessment item at DOK levels 1–3. Using this method, 17 items were selected for the content knowledge instrument (Wyse and Viger 2011).

Content validity was established using a panel of four experts, including a university faculty member in science education, a doctoral student in environmental education, a school district curriculum and professional development science specialist, and a Biology II Advanced Placement/International Baccalaureate science teacher. The experts reviewed the items for content and alignment with the lesson. Through a cycle that included review, revisions, and review, questions were modified until consensus on the quality of the items was met. The reliability of the content knowledge instrument was satisfactory ($\alpha = .724$).

Results

The study consisted of three phases. In the first phase, participants completed the pretest content knowledge instrument. This was conducted during week fifteen of the semester (1 week before the experimental phase). During the experimental phase, participants completed the 5-DIE lesson. This phase began during week sixteen of the semester. The amount of time allotted to complete the lesson was five 50-min class periods. Immediately after completing the fifth 50-min class period, participants completed the posttest content knowledge instrument. The final phase took place 5 weeks after the experimental phase when participants completed the delayed posttest content knowledge instrument.

Integrity of the Interventions

The degree to which the participants engaged in the interventions was determined by a blind review of the Research Briefs and a follow-up conversation with the teachers responsible for implementation. Ninety percent of the participants assigned to the control condition engaged in the Research Brief and 89 % of the participants in the self-explanation prompt condition engaged in the Research Brief activities as well as the self-explanation prompt scaffolding strategy, while 78 % of the participants in the faded worked example condition and 77 % of the participants in the combined condition engaged in the Research Brief activities including the self-explanation prompts and faded worked example scaffolding strategies.

A follow-up interview with the teachers indicated that the participants required little to no additional encouragement to complete the assignment. Engaged behavior was reported as being very high, with minimal need for teacher help or management. For those few participants who did struggle, teachers reported working with them one-on-one and having the participants record their ideas in their Research Brief.

Effect of the Interventions

The collective results for all participants ($N = 245$) are presented followed by the presentation of the results for the conditions. The three-factor mixed model ANOVA yielded no significant difference for the combined condition over time, $F_{(1,99)} = 2.23$ [MSE = 8.63], $p > .05$, $\eta^2 = 0.01$, no significant difference for the self-explanation condition over time, $F_{(1,43)} = .236$ [MSE = .913], $p > .05$, $\eta^2 = 0.003$, and no significant difference for the faded worked example condition over time, $F_{(1,60)} = .99$ [MSE = 3.83], $p > .05$, $\eta^2 = 0.008$. In addition, the statistical test yielded no significant difference for the combined condition, $F_{(1,241)} = 1.08$ [MSE = 22.37], $p > .05$, $\eta^2 = 0.002$, no significant difference for the self-explanation condition, $F_{(1,241)} = .010$ [MSE = .199], $p > .05$, $\eta^2 = 0.00002$, and no significant difference for the faded worked example condition, $F_{(1,241)} = 2.189$ [MSE = 45.18], $p > .05$, $\eta^2 = 0.005$. In all cases, the effect size was very small, if not negligible. The statistical test of homogeneity of variance indicated that variances in each condition were equal (Table 2).

The statistical analysis for the pretest, posttest, and delayed posttest yielded a significant difference, $F_{(1,241)} = 98.1$ [MSE = 3.80], $p < .05$, $\eta^2 = 0.17$ with a small effect size. A post hoc Tukey HSD follow-up revealed that participants performed the best overall on the posttest, indicating an acquisition of content knowledge. The delayed posttest scores were significantly lower than the posttest scores yet significantly higher than the pretest scores, indicating retention of content knowledge.

The mean score for all conditions on the content knowledge instrument between pretest and posttest increased by almost three questions (mean score difference = 2.67), while the difference in the mean score from posttest to delayed posttest decreased by less than one question (mean score difference = 0.89), and the difference in the mean score from pretest to delayed posttest increased by close to two questions (mean score difference = 1.78) (Fig. 4).

Table 2 Descriptive statistics for each level of the between factors

	Faded worked examples					
	Without			With		
	N	M	SD	N	M	SD
Pretest						
Self-explanation prompt						
Without	40	5.88	2.28	44	5.41	2.78
With	61	5.77	2.22	100	5.24	2.61
Posttest						
Self-explanation prompt						
Without	40	8.63	2.97	44	7.30	3.39
With	61	8.23	2.83	100	8.32	3.41
Delayed posttest						
Self-explanation prompt						
Without	40	7.85	3.45	44	7.05	3.35
With	61	7.23	3.64	100	7.20	3.31

Discussion

Previous empirical research indicates that learners struggle to understand and interpret scientific inscriptions (e.g., Bowen and Roth 2002) and that they struggle to learn science content while engaging in scientific argumentation (e.g., Cross et al. 2008). As a response to these learning issues, scaffolding in the form of self-explanation prompts and faded worked examples are strategies that have been shown to be successful (Atkinson et al. 2003; Chi et al. 1994). The explanation for this success has been a theorized reduction in the extraneous cognitive load associated with unguided activities, thereby shifting the intended content learning toward germane load or the cognitive processing necessary for learning (Kirschner et al. 2006). For this study, it was hypothesized that these scaffolding strategies would individually and collectively (a) shift cognitive load from extraneous to germane and (b) make explicit the science concepts, thus leading to a change in focus emphasizing the science content of the learning activity, resulting in the acquisition and retention of content knowledge. However, the statistical results indicate that the hypothesized outcomes were not achieved; learning was not improved for any condition. Yet, there was a significant small positive effect on the acquisition and retention of content knowledge across all conditions, indicating that the general learning environments was successful.

Three factors may have contributed to these results. First, since all conditions—including the control—had some form of prompting (general and/or explicit), the lesson representing the control condition was effective without the additional scaffolding strategies. Second, although the scaffolding strategies were designed to focus learners' cognitive processing on the science content represented in scientific inscriptions and in the scientific arguments, the self-explanation prompts and faded worked examples may have shifted the self-regulatory skills of self-explanation and the skill of argumentation from extraneous to germane cognitive load instead of shifting the science content to germane cognitive load. Third, the assumption of generalizability for the empirical research supporting self-explanation prompts and faded worked examples is challenged by the results of this study. Each of

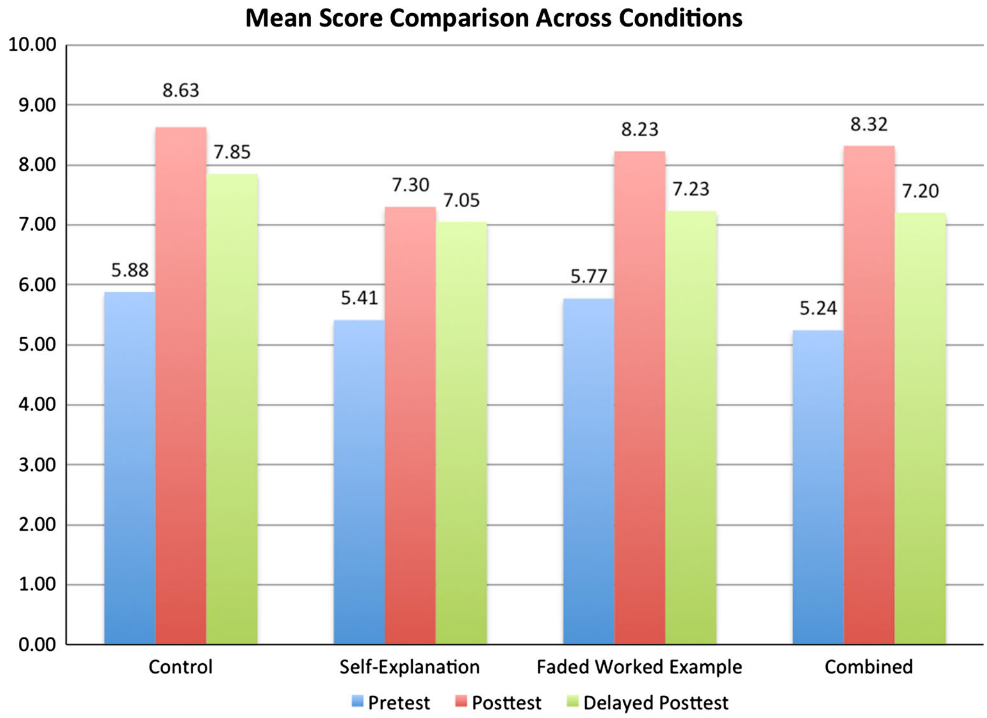


Fig. 4 Mean score for between factors. This figure displays the mean score for the pretest, posttest, and delayed posttest for each of the conditions

these factors is explained in more depth in the following sections along with an assessment of the framework for achieving its pragmatic purpose of using technology to improve the accessibility of the learning environment.

The design framework for the lesson is infused with general prompts—activity, self-monitoring, and self-explanation—and all conditions in the study, including the control, contained each of these prompts. The self-explanation condition contained both general self-explanation prompts and explicit self-explanation prompts related to the inscriptions. The faded worked example condition contained both faded worked examples and general activity prompts designed to facilitate the development of a claim statement. The combined condition contained the self-explanation and faded worked example scaffolding strategies as well as the general prompts. The control condition still contained the general prompts.

The equivalent outcomes across conditions provide evidence that learning with lessons containing general prompts was not enhanced by the addition of self-explanation prompts or faded worked examples. This conclusion aligns with the findings of van der Meij and de Jong (2011), who reported a weak effect ($\eta^2 = 0.025$) for high school physics students in an experimental session when comparing general prompts to specific prompts. Also consistent with the findings of current study, Schwonke et al. (2009), with eighth grade students in a 90-min experimental session, found no effect ($\eta^2 = 0.002$) when comparing an established cognitive tutor to the cognitive tutor enhanced with faded worked examples. van der Meij and de Jong (2011) and Schwonke et al. (2009) concluded that there was no increase in content knowledge upon addition of the extra scaffolding strategies because the control learning environments were sufficiently effective. If this is the case, it also means that metacognitive skills related to self-explanation and argumentation may be taught alongside science content with no adverse effect on the acquisition and retention of content knowledge.

The foundational empirical studies that were used to inform the design of the scaffolding strategies for this study were conducted with an older population of participants (i.e., college undergraduates and financial service employees), which may exhibit key differences when compared to the participants in this study, (Crippen and Earl 2007; Berthold et al. 2009; Kissane et al. 2008; Renkl et al. 2004). Crippen and Earl (2007) study, which demonstrated a weak effect ($\eta^2 = 0.08$), was conducted with introductory college chemistry students, all of whom were science majors. The study by Berthold et al. (2009) that demonstrated a large effect ($\eta^2 = 0.89$) involved undergraduate psychology students. The participants in the study by Kissane et al. (2008) were financial services employees, all of whom were adult learners ($\eta^2 = 0.18$); in the study by Renkl et al. (2004), conducted under strict experimental conditions, participants were undergraduate psychology students ($\eta^2 = 0.15$). All of these participant groups can be identified as highly selected populations (i.e., adult learners in the work force and college students from a select university) (Mayer et al. 2005) and not equivalent to the high school science students who were participants in the current study. Therefore, generalizing these results to participants who are fourteen to 15 years of age in the naturalistic, chaotic setting of a state-mandated biology course may be inappropriate.

The use of scaffolding as a technology-enhanced support was intended to improve the accessibility of the learning environment without adding to the instructional load of the classroom teacher. With technology providing the content and instruction, this represents one form of using the affordances of cyberlearning (Martinez and Peters Burton 2013) to address the pragmatic constraints of technology integration (Keengwe et al. 2008), including the practical needs of teachers. Beyond the current study, the efficacy and details of such a strategy would have significance for applications of blended learning (i.e., mixing online and face-to-face content and instruction) as well as for self-directed online learning. In this case, with participating teachers indicating a high level of engagement, little need for support, and time and opportunity for providing one-on-one help, the design framework did achieve its purpose of providing a technology-enhanced experience for students with the teacher serving in an adaptable supporting role. These results are consistent with the previous iterations of research related to the design framework (Kern et al. 2014), but the minimal reporting of previously known issues, such as split attention due to multiple software windows or the need for embedded content area literacy strategies, also demonstrates improvements in the usability for teachers.

Limitations of the Study

Due to the quasi-experimental design of this study, participants were not randomly assigned to the treatment and control groups. This increases the risk that there may have been preexisting differences between participants across the range of courses related to test-taking variables such as reading level, test anxiety, and variation in testing conditions. During the study, several disruptions to the learning environment occurred, such as network computing at the school and district level failing several times, a fire drill, and a threat of violence against the students, which resulted in high number of absences on the day of the posttest.

The content knowledge instrument was specifically designed to prevent a ceiling effect, meaning questions were designed with the participants with the highest efficacy in the content and test taking in mind. The participant pool consisted primarily of inclusionary and general biology participants (77.5 %). The inclusionary and general biology participants included participants with documented disabilities and English language learners. This may have contributed to poor test performance due to low reading levels, test anxiety, and item difficulty.

Conclusion

This study was meant to address two educational needs related to learner difficulties with learning science in a cyberlearning environment when engaging with scientific inscriptions and argumentation. First, research-based conjectures were made that self-explanation prompts and faded worked examples would shift science content represented in inscriptions and scientific arguments from extraneous to germane load. Second, these conjectures were explored through the design, implementation, and evaluation of the lesson. The findings revealed a need for more research evaluating how theory developed outside the naturalistic setting of a K-12 science classroom translates to that context.

The research and development of the lesson format used in this study comes as a response to the need for an inquiry-based cyberlearning environment meant to promote critical thinking, problem solving, and scientific literacy. The time and effort allocated to constructing self-explanation prompts and faded worked examples for a lesson may not be necessary when the 5-DIE approach is employed because it already includes a rich variety of general scaffolds. This guidance would also apply to other design frameworks for this context; the use of general scaffolding strategies are as effective as more nuanced varieties, like those investigated in this study that are also more labor intensive to construct.

Finally, the DBR process is meant to inform the researcher's understanding of teaching and learning. The decisions made by teachers in K-12 education are informed by theories that may not be generalizable to their context. This study suggests a potential disconnect between research with the highly self-regulated world of college learners and the more generalized population of K-12 learners. As evidenced by this study and the assertions of Barab and Squire (2004) that learning, cognition, knowing, and context cannot be treated in isolation, considerations must be made for the tremendous differences in prior knowledge, self-regulatory skills, and motivation that exist between the college and pre-college learners when making K-12 curriculum design decisions. This research, as well as future research endeavors catalyzed by this study, should be centered on the naturalistic and often times chaotic context of the K-12 science classroom where countless variables influence the learning. This study reinforces the need for research where K-12 learning occurs—in the classroom.

References

Atkinson RK, Derry SJ, Renkl A, Wortham D (2000) Learning from examples: instructional principles from the worked examples research. Rev Educ Res 70(2):181–214

Atkinson RK, Renkl A, Merrill MM (2003) Transitioning from studying examples to solving problems: effects of self-explanation prompts and fading worked-out steps. J Educ Psychol 95(4):774–783. doi:10.1037/0022-0663.95.4.774

Barab S, Squire K (2004) Design-based research: putting a stake in the ground. J Learn Sci 13(1):1–14

Berthold K, Eysink THS, Renkl A (2009) Assisting self-explanation prompts are more effective than open prompts when learning with multiple representations. Instr Sci 37(4):345–363

Bowen G, Roth W (2002) Why students may not learn to interpret scientific inscriptions. Res Sci Educ 32(3):303–327

Bricker LA, Bell P (2008) Argumentation from science studies and the learning sciences and their implications for the practices of science education. Sci Educ 92(3):473–498

Chi MT, De Leeuw N, Chiu M, Lavancher C (1994) Eliciting self-explanations improves understanding. Cognit Sci 18(3):439–477. doi:10.1016/0364-0213(94)90016-7

Clark RE, Kirschner PA, Sweller J (2012) Putting students on the path to learning: the case for fully guided instruction. Am Educ 36(1):6–11

Crippen KJ, Earl BL (2007) The impact of Web-based worked examples and self-explanation on performance, problem solving, and self-efficacy. Comput Educ 49(3):809–821. doi:10.1016/j.compedu.2005.11.018

Cross D, Taasoobshirazi G, Hendricks S, Hickey DT (2008) Argumentation: a strategy for improving achievement and revealing scientific identities. Int J Sci Educ 30(6):837–861

Glazer N (2011) Challenges with graph interpretation: a review of the literature. Stud Sci Educ 47(2):183–210. doi:10.1080/03057267.2011.605307

Kalyuga S (2009) Instructional designs for the development of transferable knowledge and skills: a cognitive load perspective. Comput Hum Behav 25(2):332–338

Kalyuga S, Ayres P, Chandler P, Sweller J (2003) The expertise reversal effect. Educ Psychol 38(1):23–31

Keengwe J, Onchwari G, Wachira P (2008) Computer technology integration and student learning: barriers and promise. J Sci Educ Technol 17(6):560–565

Kern C, Crippen KJ (2013) A design framework and research program for enacting science cyberlearning. In: Bastiaens T, Marks G (eds) Proceedings of e-learn: world conference on e-learning in corporate, government, healthcare, and higher education 2013. Association for the Advancement of Computing in Education (AACE), Chesapeake, VA, pp 1195–1199

Kern C, Crippen KJ, Skaza H (2014) The fidelity and usability of 5-DIE: a design study of enacted cyberlearning. J Comput Math Sci Teach 33(1):513–536

Kim MC, Hannafin MJ (2011) Scaffolding problem solving in technology-enhanced learning environments (TELEs): bridging research and theory with practice. Comput Educ 56(2):403–417

Kirschner P, Sweller J, Clark RE (2006) Why unguided learning does not work: an analysis of the failure of discovery learning, problem-based learning, experiential learning and inquiry-based learning. Educ Psychol 41(2):75–86

Kissane M, Kalyuga S, Chandler P, Sweller J (2008) The consequences of fading instructional guidance on delayed performance: the case of financial services training. Educ Psychol 28(7):809–822

Lee H, Liu O (2010) Assessing learning progression of energy concepts across middle school grades: the knowledge integration perspective. Sci Educ 94(4):665–688

Martinez ME, Peters Burton EE (2013) Cognitive affordances of the cyberinfrastructure for science and math learning. Educ Media Int 48(1):17–26. doi:10.1080/09523987.2010.535333

Mayer RE, Hegarty M, Mayer S, Campbell J (2005) When static media promote active learning: annotated illustrations versus narrated animations in multimedia instruction. J Exp Psychol Appl 11(4):256–265

McKenney S, Reeves TC (2012) Conducting educational design research. Routledge, London

Notar CE, Zuelke DC, Wilson JD, Yunker BD (2004) The table of specifications: insuring accountability in teacher made tests. J Instr Psychol 31(2):115–129

Renkl A, Atkinson RK, Grobe CS (2004) How fading worked solution steps works: a cognitive load perspective. Instr Sci 32(1):59–82

Rew L, Becker H, Cookston J, Khosropour S, Martinez S (2003) Measuring cultural awareness in nursing students. J Nurs Educ 42(6):249–257

Roy M, Chi MT (2005) The self-explanation principle in multimedia learning. In: Mayer RE (ed) The Cambridge handbook of multimedia learning. Cambridge University Press, Cambridge, pp 271–286

Ruiz-Primo MA, Shavelson RJ, Hamilton L, Klein S (2002) On the evaluation of systemic science education reform: searching for instructional sensitivity. J Res Sci Teach 39(5):369–393

Schnotz W, Kürschner C (2007) A reconsideration of cognitive load theory. Educ Psychol Rev 19(4):469–508

Schwonke R, Renkl A, Krieg C, Wittwer J, Aleven V, Salden R (2009) The worked-example effect: not an artifact of lousy control conditions. Comput Hum Behav 25(2):258–266

Staker H, Horn MB (2012) Classifying k-12 blended learning. http://www.innosightinstitute.org/innosight/wp-content/uploads/2012/05/Classifying-K-12-blended-learning2.pdf

Sweller J (1988) Cognitive load during problem solving: effects on learning. Cogn Sci 12(2):257–285

Sweller J, Cooper GA (1985) The use of worked examples as a substitute for problem solving in learning algebra. Cogn Instr 2(1):59–89

van der Meij JJ, de Jong TT (2011) The effects of directive self-explanation prompts to support active processing of multiple representations in a simulation-based learning environment. J Comput Assist Learn 27(5):411–423

van Gog T, Paas F, van Merrienboer JG (2008) Effects of studying sequences of process-oriented and product-oriented worked examples on troubleshooting transfer efficiency. Learn Instr 18(3):211–222

van Merriënboer J, Kirschner PA, Kester L (2003) Taking the load off a learner's mind: instructional design for complex learning. Educ Psychol 38(1):5–13

von Aufschnaiter C, Erduran S, Osborne J, Simon S (2008) Arguing to learn and learning to argue: case studies of how students' argumentation relates to their scientific knowledge. J Res Sci Teach 45(1):101–131

Webb NL (2007) Issues related to judging the alignment of curriculum standards and assessments. Appl Meas Educ 20(1):7–25

Wyse AE, Viger SG (2011) How item writers understand depth of knowledge. Educ Assess 16(4):185–206