Self-Efficacy Change Associated with a Cognitive Load-Based Intervention in an Undergraduate Biology Course

David F. Feldon, Joana Franco
Utah State University

Jie Chao
Concord Consortium

James Peugh
Cincinnati Children's Hospital and Medical Center

Cathy Maahs-Fladung
Utah State University

April 20, 2017

Self-Efficacy Change Associated with a Cognitive Load-Based Intervention in an Undergraduate Biology Course

February 26, 2018

Abstract

Cognitive load theory (CLT) holds that discovery learning and other instructional strategies imposing high levels of extraneous load on novice learners hinder learning. Such learning conditions are also associated with significant drops in persistence, a key measure of motivation. However, research within the CLT framework typically engages motivation as a necessary precursor to learning, rather than as an outcome of instruction. In this study, we examine changes in motivational beliefs as outcomes of learners' cognitive processes through a CLT lens as they engage with instruction. Using a double-blind quasi-experimental design, we manipulate the level of cognitive load imposed on participants through instruction and assess changes in self-efficacy from pre- to post-intervention. In an analysis of data from students enrolled in an undergraduate biology course (n=2,078), students in the treatment condition demonstrated significantly higher performance on end-of-semester lab reports and self-efficacy measures. However, post-instruction self-efficacy was not significantly related to performance, controlling for pre-instruction self-efficacy, gender, and scientific reasoning ability. These findings introduce the possibility that the cognitive load imposed on working memory during instruction may affect motivational beliefs and provides a foundation to further explore connections between historically distinct theoretical frameworks such as CLT and social cognitive theory.

Keywords: cognitive load, motivation, self-efficacy, mental effort, cognitive task analysis.

Changing Self-Efficacy as a Function of Manipulating Cognitive Load in an Undergraduate Biology Course

In cognitive load theory (CLT), motivational beliefs are considered primarily to be a precursor, rather than an outcome, of instruction (Moreno & Mayer, 2007). CLT research studies typically assume that sufficient motivation is required for participants to invest the mental effort necessary to meet the cognitive demands of instruction (van Merriënboer & Sweller, 2005; Kanfer & Ackerman, 1989). However, nascent work has begun to consider further the nature of the relationship between learning as a function of CLT-based instructional principles and the role of motivation (e.g., Paas, Tuovinen, van Merriënboer, & Darabi, 2005; Likourezos & Kalyuga, 2017; Schnotz, Fries, & Horz, 2009; van Gog & Rummel, 2010). In these analyses, invested mental effort is considered a nexus between cognitive and motivational perspectives as an index of both imposed cognitive load (assuming motivation sufficient to engage for the duration of the learning task; Paas, 1992) and motivation (Pintrich, 1990; Schunk, Pintrich, & Meece, 1996; Wigfield & Eccles, 2000). For example, the imposition of excessive cognitive load is associated with drops in persistence, which is operationally defined as sustained mental effort until the completion of a goal (Britt, 2005; Lewis, Bishay, McArthur, & Chou, 1993; Paas et al., 2005). However, Schnotz and colleagues speculate that stripping too many interesting-but-extraneous details from instruction may result in learning materials that are “no longer optimally activating from a motivational perspective” (p. 81) and consequently decrease invested effort. We test the hypothesis that efficient management of cognitive load can result in positive shifts in measures of motivational belief. The findings of the study presented here suggest that motivational belief (i.e., self-efficacy) may be a consequence of the cognitive load imposed by instruction, rather than merely a necessary precursor of the decision to invest mental effort.

Mental Effort in the Context of Cognitive Load Theory

From the perspective of cognitive load theory, the major factor influencing an individual's success in learning from instruction is the limited ability of working memory to assimilate and structure target information. Working memory capacity is generally considered to be capable of processing very few pieces of information at a time and of retaining them for less than 20-30 seconds without rehearsal (Cowan, 2001; van Merriënboer & Sweller, 2005). In that sense, working memory functions as a bottleneck, filtering the information to be encoded in long-term memory through attentional, conscious processes in ways that are evolutionarily adaptive for information processing (Sweller, 2004). The availability of relevant and well-structured prior knowledge increases the functional capacity of working memory relative to the task, such that an individual with greater expertise will experience a lower burden on working memory resources than an individual with less expertise (Ericsson & Kintsch, 1995; Author, 2007; Gobet, 1998; Sweller, 1994).

The capacity of working memory can be operationally defined by the maximum quantity of new, non-automated information it is capable of processing at a given time. As a corollary, the greater the quantity of non-automated or novel information to be processed (i.e., cognitive load), the greater the requirement to invest mental effort for successful processing (Kalyuga, 2011). When the cognitive load imposed exceeds the working memory capacity of the learner, maximal investment of mental effort on the part of the student will not be sufficient to attain the intended learning or performance outcomes (Paas & van Merriënboer, 1993).

In dealing with difficult tasks, higher degrees of cognitive demand impose higher load and require greater effort. In other words, “mental load is imposed by instructional parameters (e.g., task structure, sequence of information), and mental effort refers to the amount of capacity that is allocated to the instructional demands” (Paas, 1992, p. 429). If cognitive load imposed by instructional material exceeds the level of effort an individual can or does invest, instruction will be less effective than if effort is greater than or equal to the demands imposed by the learning task. Learning tasks that have been practiced consistently require less conscious information processing in working memory due to the development of automated knowledge (i.e., learned, unconscious processing) (Anderson, 1982; Blessing & Anderson, 1996; Clark, 2014) and schema development (van Merrienboer & Sweller, 2005).

Types of Cognitive Load

CLT currently identifies three categories of cognitive load that might be imposed on a learner during the learning process: intrinsic, extraneous, and germane (van Merriënboer & Ayres, 2005; Kalyuga, 2011). For effective learning to occur, the sum of these loads must remain smaller than the capacity of the learner’s working memory. Therefore, the main objective of CLT has been to derive principles for managing cognitive load during instruction to maximize the efficiency and effectiveness of instruction (Paas et al., 2003; Tuovinen & Paas, 2004).

As originally established by Sweller (1993, 1994), intrinsic cognitive load is a characteristic of the information to be learned itself, independent of the learner. Thus, information that entails more propositions or more interactions among knowledge elements imposes a higher level of intrinsic load by definition (van Merriënboer & Sweller, 2005). More recent studies, however, argue that the level of intrinsic load is also influenced by “the degree of interactivity between essential elements of information relative to the level of learner expertise in the domain" (Kalyuga, 2011, p. 2). As such, an individual with higher levels of relevant and accurate prior knowledge will process information with a lower burden on working memory (i.e. intrinsic cognitive load) than an individual with a lower level of prior knowledge. Further, this approach permits convergence between the intrinsic and germane load constructs, with the total quantity of cognitive load necessary for optimal learning represented by the learner's capacity for processing the instructional content itself combined with the appropriate instructional mechanisms necessary for optimal learning to take place.

Extraneous cognitive load is imposed by burdening working memory during instruction in a manner that does not positively contribute to learning. This type of load is associated with inappropriate instructional design and activities, which can manifest in two possible ways. First, instruction or instructional materials may force a learner to process unnecessary or irrelevant information that results in unproductive element interactivity in working memory (Ayres & Paas, 2012; Kalyuga, Chandler, & Sweller, 1999; Mayer, Heiser, & Lonn, 2001; Sweller, van Merriënboer, & Paas, 1998, Sweller, 2010). Second, information necessary or beneficial to instruction may be withheld, which forces a learner to simultaneously structure and attempt to solve a problem for which appropriate schemas are not yet developed (Likourezos & Kalyuga, 2016; Sweller, 1988). Similarly, "any instructional procedure that requires learners to engage in… a search for referents in an explanation (i.e., when Part A of an explanation refers to Part B without clearly indicating where Part B is to be found) is likely to impose a heavy extraneous cognitive load because working memory resources must be used for activities that are irrelevant to schema acquisition and automation" (Paas, Renkl, & Sweller, 2003, p. 2). Thus, when guidance is needed and not provided, cognitive information processing becomes a burden to learners and likely ineffective for learning (Kirschner, Sweller, & Clark, 2006).

Mental Effort in the Context of Motivation Theories

Theories of motivation consider investment of mental effort to be one of three major indicators of motivation, along with goal selection (a decision of where to invest mental effort) and persistence (the maintenance of mental effort over time until a goal is achieved) (Pintrich, 1990; Schunk et al., 1996; Wigfield & Eccles, 2000). When motivated, learners also tend to demonstrate a more strategic approach to learning tasks and direct mental effort toward processes that are more pertinent to learning (Rey & Buchwald, 2011).

One of the most prominent theories of motivation that links beliefs to effort investment is social cognitive theory (SCT; Bandura, 1992, 1997). SCT holds that self-efficacy (i.e. one's belief in their capability to manage and succeed in a particular task) drives the investment of mental effort (Bandura, 1992), because "unless people believe that they can produce desired effects by their actions, they have little incentive to act" (Bandura, Barbaranelli, Caprara, & Pastorelli, 1996, p. 1206). Further, successful past performances can enhance self-efficacy, contributing to higher goal aspirations and further investment of effort, which produce subsequent performance improvements (Bandura, 1997). This perspective has driven a large proportion of motivation studies in education, with meta-analyses supporting the positive relationship between self-efficacy and achievement (Bandura & Locke, 2003; Multon, Brown, & Lent, 1991; Stajkovic & Luthans, 1998).

Anticipated Investment of Mental Effort

From the SCT perspective, learners' beliefs about the necessary level of effort to invest in a learning task is of central importance. Similarly, CLT assumes that in order for instruction to be effective, students need to be motivated so that they will invest sufficient mental effort to meet the cognitive demands imposed by the instruction (van Merriënboer & Sweller, 2005). Such motivation is typically indicated by the learner's choice to engage in a given learning task (i.e. goal selection), so that if the perception of task difficulty is extremely high, it could lead to a lack of engagement (Clark, 1999). Salomon (1984) argued that students "make judgments on the basis of the perceived attributes of the instructional procedures, and subsequently expend mental effort accordingly" (p. 649).

For example, Zheng, McAlack, Wilmes, Kohler-Evans, & Williamson (2009) found that participants receiving instruction within an interactive multimedia context reported greater self-efficacy than their counterparts in a non-interactive version. In this case, the participants' perceptions and expectations regarding instructional format were highly salient, because self-efficacy mediated the relationship between instructional condition and task performance. While the influence of instructional condition on self-efficacy and the influence of self-efficacy on performance were each positive, the direct effect of instructional condition on performance was negative, indicating the importance of motivational beliefs for influencing learning outcomes even when instructional design may hinder learning outcomes.

As individuals develop knowledge and skills that support successful task performance, their perceptions of the necessary effort required to perform the task decreases, resulting in diminishing estimates of necessary effort and subsequent allocation of effort to performance (Yeo & Neal, 2008). Similarly, it is possible that changes in belief regarding necessary effort may stem from perceptions of effort expended during learning and be expressed by participants in the form of self-efficacy beliefs (Clark, 1999). For example, in a study comparing constructive failure and direct instruction strategies, Likourezos and Kalyuga (2017) found that participants' perceived mental effort, perceived task difficulty, and expected probability of task success (i.e., self-efficacy) differed significantly as a function of instruction. Participants receiving fully guided instruction with worked examples reported significantly lower levels of perceived effort and task difficulty and higher levels of self-efficacy than the unguided problem-solving condition representing the constructive failure approach. However, these differences were obtained in the absence of significant differences across instructional conditions in posttest performance. The authors concluded that differences in learners' goals and possible low levels of task complexity could account for the results. However, it is also possible that cognitive load can have a direct impact on motivational beliefs, even in the absence of differences in learning outcomes.

Research Questions

In the current study, we argue that these learners' estimates of necessary effort, and thus their self-efficacy, may be influenced directly by the level of cognitive load imposed by instruction. As such, motivational beliefs can be outcomes rather than merely predictors of the learners' cognitive processes as they engage with instruction. Specifically, the level of extraneous cognitive load imposed during instruction may be associated with changes in learners' expectations regarding the necessary levels of effort required in future, related tasks, independent of performance levels. Thus, we address the following research question: Can a manipulation of cognitive load undetected by participants predict a differential change in post-instruction motivational beliefs? We hypothesize that:

1. Participants in the treatment (lower extraneous load) condition will demonstrate stronger

performance in the post-instruction assessment than participants in the control (higher extraneous load) condition.

2. Participants in the treatment condition will demonstrate greater gains in self-efficacy from pre- to post-instruction than participants in the control condition.
3. Post-instruction self-efficacy levels cannot be accounted for by differences in task performance, reflecting the influence of cognitive load imposed by instruction rather than beliefs formed on the basis of assessment task performance.

Method

Participants

Participants in this study were undergraduates at a public research university in the Southeastern United States who were enrolled in a one-semester introductory biology course that was offered every Fall and Spring term. Data for this study were drawn from 5 consecutive terms from Spring, 2008 through Spring, 2010 (N=2,078; n = 1,052 treatment, n = 1,026 control). The course consisted of 3 lecture hours and 3 laboratory hours per week, providing a survey of macromolecules, cell structure and function, genetics, and molecular biology. As the course primarily served freshman in biology and allied health majors, the course material was generally relevant to their future goals.

Thirty-eight percent of the participants were male, and 62% were female. The average age was 19.6 years. Seventy-eight percent of the participants majored in biology-related disciplines such as biology, biomedical engineering, nursing, pharmacy, and exercise science; 22% majored in typically unrelated disciplines such as computer science, economics, art, history, etc. All participants were blind to experimental conditions and to the existence of the study, as were the graduate teaching assistants who taught the weekly laboratory sections. Participants did not need to provide consent for data collection, because the study was granted exempt status by the university's institutional review board. Its activities occurred as part of normal educational practice using instruments typical of the university classroom environment.

Materials

In addition to weekly lectures and laboratory sessions, students were required to watch brief weekly videos (~10 minutes) that explicitly instructed students in the processes of biological research, beginning with the identification of potentially productive trends in observed data or primary literature, the framing of research questions and testable hypotheses, the design of experiments, the analysis of data, and the drawing of justifiable conclusions based on data. Two versions of the videos were created. The first were recordings of a biology professor at the university, experienced with teaching the introductory biology course, who had won multiple teaching awards delivering his own brief lectures on these topics. The second were recordings of the same professor in the same setting delivering lectures on the same topics. However, in the second (i.e., treatment) condition, the script was provided to the professor and was developed on the basis of cognitive task analyses (CTA) (See Author et al., 2010). Thus, the differences between the two sets of videos did not lie in what content was covered. Instead, they differed in the level of detail provided and organizational structure provided in each lecture.

CTA uses interviewing techniques and other knowledge elicitation methods to capture both explicit and tacit knowledge from experts (Author et al., 2008). Because experts' procedural knowledge tends to be automated and their schemas tend to be highly efficient in the organization of relevant information, they frequently but unintentionally omit information on how they solve problems in their domains of expertise (Clark, 2009; Clark et al., 2012; Author, 2007, 2010; Rikers, Schmidt, & Boshuizen, 2000; Sullivan, Yates, Inaba, Lam, & Clark, 2014). Consequently, CTA-based training is typically more complete, providing a higher level of detail than instructional content identified through other means (Author et al., 2009). CTA-based training also typically yields greater learning gains than training based on other sources of instructional content. A recent meta-analysis found a large advantage for CTA-based training, reporting a Hedge's *g* mean effect size of 0.87 (Author et al., 2013).

As discussed previously, cognitive load theory predicts that incomplete instruction or instruction omitting important information for learners (e.g., detailed step-by-step processes) imposes extraneous load that hinders student learning (de Jong, 2010; Mayer & Moreno, 2003). For a novice learning the complex task of science inquiry—as in this study—the limitations of working memory are accentuated due to the combination of both the great amount of new content information and the complexity of the task (i.e., problem solving; Sweller, 1988) in the absence of well-structured schemas that could retain relevant information without additional effort. As such, the effectiveness of CTA-based training that scaffolds procedural learning is typically attributed to a reduction in extraneous load imposed by otherwise missing information regarding detailed steps and rules that is captured through the CTA process (Author et al., 2010). To produce the instructional videos, the lead author conducted CTA interviews with three experts in biological research salient to the focus on the course. Details regarding the CTA procedures and outcomes used to develop the instructional videos for this study are reported in detail elsewhere (Author et al., 2010; Author et al., 2009).

To prevent the CTA-based scripts from inadvertently influencing the recorded business-as-usual lectures, all control condition videos were created before the professor was provided with the CTA-based scripts. Eight pairs of videos (traditional versus CTA-based) were developed. Each video lasted 5-10 minutes. Content analysis of the videos indicated that the explanations provided in the traditional videos were more abstract and presented principles illustrated with examples (Author et al., 2009). In contrast, CTA-based videos provided more specific and detailed statements and were framed as a set of step-by-step actions and decisions to be made. To maintain the students' viewing habit every week, two condition-neutral videos were also developed to match video deployment dates to each week's scheduled content.

Comparison of cognitive load imposed by treatment and control videos. Treatment and control versions of the videos were viewed under laboratory conditions by undergraduate participants who did not enroll in the biology course (n=42). These participants rated the cognitive load imposed by the full set of videos they viewed (either treatment or control) using Paas's (1992) 9-point Likert item. Mean cognitive load ratings indicated that the treatment condition imposed less total load than the control ($Mean_{treatment} = 5.35$, $SD_{treatment} = 2.01$; $Mean_{control} = 5.56$, $SD_{control} = 1.46$), though the difference was not statistically significant.

As reported by Author et al. (2009), the treatment videos presented higher levels of instructionally relevant information (i.e., information that would impose intrinsic cognitive load) in 39.8% of coded transcript segments (treatment = 33 out of 83; control = 0 out of 83) with an interrater agreement rate of 97% (disagreements resolved through discussion). Because total cognitive load experienced by a learner is the sum of extraneous cognitive load and intrinsic cognitive load processed in working memory (i.e., cognitive load is "additive"; van Merrienboer & Sweller, 2005, p. 150), a difference in intrinsic load between conditions without a concomitant difference in overall cognitive load must, by definition, reflect differential levels of extraneous load. Considering both the lower level of overall cognitive load reported by participants viewing the treatment materials and the higher level of intrinsic load identified in the treatment condition through content analysis, we conclude that the treatment condition must have imposed less extraneous load than the control condition.

A form of cognitive efficiency was computed on the basis of the perceived level of cognitive load relative to quantity of assessment-relevant information (i.e., intrinsic load) included in the videos of each condition. Dividing the mean levels of perceived load by the number of instances of greater intrinsic load for each condition (increasing the count for the control condition to 1 to avoid dividing by 0) yielded a ratio of 0.162 for the treatment condition and 5.56 for the control condition. A two-sided Z-test confirmed a significant difference indicating that cognitive efficiency (perceived load/intrinsic load) was significantly greater in the treatment condition ($Z = 8.4$, $p < 0.001$).

Course structure. Following the completion of the series of videos, the laboratory portion of the course culminated in a multiple-week, inquiry-based investigation of *Drosophila melanogaster* (fruit fly) genetics wherein students were required to make observations, generate hypotheses, collect, analyze and interpret data and form conclusions based on those data. The work product submitted for course credit was a formal paper, written in scientific format, reporting their findings. Due to logistical constraints, all students investigated the same unknown genetic cross. Their task was to determine the genotypes of the parental generation and if the alleles exhibited Mendelian inheritance patterns. Investigations were conducted in small groups within a laboratory section. *Drosophila* observation data were pooled within each laboratory section to increase sample size. Thus, students were provided with the research question, hypothesis, and methods, but they had complete discretion in the scientific judgments articulated in their papers (i.e., discussions of intellectual context, study rationale, data analysis and interpretation, conclusions, and limitations). Lab reports were written and submitted individually by students via an online course management website and scored using a rubric described in the *Measures* section below. All papers were checked for plagiarism using *SafeAssign*™ and papers containing plagiarized material were not included in the sample.

Procedure

This study employed a double-blind, quasi-experimental design with random assignment to conditions to evaluate the impact of CTA-based instruction on undergraduate biology students' motivation and achievement. We randomly assigned laboratory sections to either use CTA-based online instructional videos (experimental condition) or traditional online instructional videos (control condition). Laboratory sections were facilitated by graduate teaching assistants (TAs) who were assigned to a single condition (i.e., no TA taught sections in both the treatment and control conditions). Both sets of videos were similar except the actual verbiage of the content being presented (e.g., same presenter, same clothing, same room, same topics, etc.). The participants and TAs were blind to the existence of the study and the instructors and researchers were blind to participant assignment to experimental conditions, which averted both potential experimenter effects and the Hawthorne effect (Rosenthal, 1966).

Students' viewing behavior was recorded via server logs, and points contributing to the final course grade were awarded for viewing the videos outside of class each week. Viewing rates ranged from 69% to 93% of students each week ($mean_{control}$ = 87% of students; $mean_{treatment}$ = 82% of students). Viewing was consistently but non-significantly higher in the control condition each week, with only 1 of the 12 videos being viewed more in the treatment condition (93% compared to 92%). As students watched the videos in laboratory each week as well, these viewings represent reinforcement of the material. Thus, there do not appear to be any meaningful distinctions between viewing rates in treatment and control and if such differences exist, they favor the control condition.

To ensure the equivalency of students' general scientific reasoning ability and initial levels of motivation between conditions, Lawson's Test of Scientific Reasoning (Lawson, 1978, 2000) and two subscales of the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich, Smith, Garcia, & McKeachie, 1991, 1993) were both administered at the beginning of the course in every semester. During the course in every semester, the traditional and CTA-based instructional videos were delivered via the internet as a series of streaming videos as described above. In both control and experimental condition, each condition-appropriate video became available for viewing at the beginning of the week for which it was assigned and students were required to view it before each weekly laboratory session (students were free to view it as many times as they wished). Viewing the videos was a required weekly assignment for students, for which they received a small number of points toward their final grade in the course. At the beginning of each laboratory session, students viewed the videos again with their TA and briefly discussed the content of videos. The MSLQ was administered in class for a second time, just prior to the end of the course each semester. Learning outcomes were assessed using the final written lab report of the semester, in which students were provided with a testable hypothesis and methods and the asked to write a report of findings that emphasized primarily the discussion section against the rubric criteria listed in the preceding section.

Measures

To assess participants' motivation, two subscales from the MSLQ (Pintrich et al., 1991) served as dependent variable measures in this study. The MSLQ is a self-report instrument to measure undergraduates' motivation and learning strategies specific to a designated course with strong reliability and validity ($\alpha \geq 0.90$ for undergraduate populations; Spitzer, 2000). The current study utilized the task value (6 items) and self-efficacy (8 items) subscales. In this case, the instrument was situated specifically within the laboratory section of the biology course, which had an explicit focus on development of authentic biology research skills and appropriate interpretation of empirical results. As such, the subscales were appropriate for eliciting participants' motivational beliefs in a context with known "situational demands" (Bandura, 2007, p. 646). Confirmatory factor analysis conducted with the pretests from the sample in the current study replicated the factor structures reported by Pintrich et al. (1993): each of the relevant items loaded onto only the anticipated factor with weights between 0.755 and 0.907. Attained reliability for the task value and self-efficacy scales using the current sample were $\alpha = 0.924$ and $\alpha = 0.946$, respectively.

Lawson's Test for Scientific Reasoning (Lawson, 1978, 2000) assesses participants' abilities to distinguish between discrete sources of variance, apply proportional, probability, and correlation reasoning, control of variables, use hypothetical-deductive method. It consists of 24 multiple-choice items and has been validated with high reliability ($\alpha = 0.81$; Lawson, Alkhoury, Benford, Clark, & Falconer, 2000) among undergraduates. On the test items, participants are presented with several scenarios and required to draw correct deductive inferences from presented data and evaluate the effectiveness of strategies to control variables. Attained reliability for the sample in the current study was $\alpha = 0.865$.

As described in Author et al. (2010), student learning gains were measured using rubric-based assessment of sole-authored lab reports. Development and validation data are presented in Timmerman, Strickland, Johnson, & Payne (2011). The rubrics measured the quality of demonstrable skills in the research process that were evident in the lab reports, specifically: the ability to set the research in context, cite relevant literature and concepts accurately, generate testable hypotheses, generate hypotheses with scientific merit, appropriately select data, effectively present data, appropriately analyze data, base valid conclusions on data, generate and evaluate alternative explanations for results, identify the limitations of the study design, and generate implications and gauge significance of the findings.

Overall reliability of the rubric was calculated using generalizability analysis and found to be high ($g = 0.85$). Attained pairwise inter-rater reliability for the current sample ranged from $g = 0.70$ to 0.86 for each rubric plank (Timmerman et al., 2011). Analysis of the factor structure using Mplus (Version 7.4) for the rubric yielded four factors, described as Framing the Study, Hypotheses, Results, and Discussion. Framing the Study consisted of two rubric planks: (1) Setting the work in context (theoretical importance) ($b = 0.75$) and (2) Accuracy and relevance of information cited ($b = 1.00$). Hypotheses also consisted of two planks: (1) Testability of hypotheses ($b = 0.84$) and (2) Scientific merit of hypotheses ($b = 1.00$). Results consisted of three planks: (1) Data selection ($b = 0.82$), (2) Data presentation ($b = 0.27$), and (3) Data analysis ($b = 1.00$). Discussion consisted of four planks: (1) Conclusions based on data ($b = 0.55$), (2) Alternative explanations for data ($b = 1.35$), (3) Limitations of study design ($b = 1.01$), and (4) Implications/significance of research ($b = 1.00$). Planks related to research methods were excluded from all analyses, because the methodology (i.e., experimental design, measures, etc.) were determined by the instructor, as described previously. Factors intercorrelated minimally but significantly ($0.04 \leq r \leq 0.10$), and overall model fit was good ($X^2 = 304.75$, df=36, $p < .001$; RMSEA = 0.073; CFI = 0.934; SRMR = 0.046). Primary evaluative weight was placed on the Discussion scores, due to the anticipated similarity of scores on Hypotheses and Results due to the common methodology dictated by the instructor and used across all laboratory sections in both treatment groups (see Author et al., 2010).

Data Analysis

Tests of the three hypotheses utilized a multivariate analysis of covariance (MANCOVA) framework within a hierarchical linear model to control for nested variance at the laboratory section level using specific commands ('Type = Complex') in Mplus (Version 7.4) that allow the ignoring of nesting without producing biased parameter estimates. All group comparison analyses were conducted using the multiple-group analysis function in Mplus to ensure that the MANCOVA assumption of homogeneity of covariate regression slopes is met through parameter estimate constraints while appropriately handling missing data. Across all variables per case, missing data ranged between 21.1% - 41.3%, with no individual variable exceeding 32.7% missing (including participants who dropped or withdrew from the course or failed to turn in assignments). Missing data was handled via the default maximum likelihood estimation algorithm in Mplus (MLR).

Analyses were conducted in two phases. In the first phase, group differences on the rubric factor scores (i.e., observed means of items loading significantly onto each factor), as well as post-test self-efficacy and post-test task value scores were tested, using gender, Lawson pretest scores, and self-efficacy pre-test and task value pre-test scores, respectively, as covariates. In the second phase, group differences on the self-efficacy and task value were tested using lab report factor scores as covariates to control for the potential effect of performance on post-test self-efficacy and task value. Statistical significance was determined using two-tailed p-values, and effect sizes were calculated using Cohen's *d*.

Results

Performance Outcomes

Tests of the first hypothesis yielded a statistically significant group difference on one of the four factor scores on the laboratory report assessment: treatment group participants showed stronger performance on their Discussion scores (mean difference = 0.45; $p < .05$; $d = 0.36$; see Figure 1), after controlling for gender and Lawson's Test scores (see Table 1 for unique R^2 of covariates). As discussed above and elsewhere (Author et al., 2010), performance on the Discussion factor was considered the most targeted measure of efficacy for this intervention. This anticipated difference between conditions supports the assumption that the CTA-based instruction decreased extraneous load compared to the control condition.

[INSERT FIGURE 1 ABOUT HERE]

[INSERT TABLE 1 ABOUT HERE]

Self-Efficacy and Task Value

Tests of the second hypothesis yielded a marginal group difference: treatment group participants showed greater post-instruction self-efficacy scores (mean difference = 3.90; p = 0.093; $d = 0.54$; see Figure 2), after controlling for pretest self-efficacy scores, gender, and Lawson's Test scores (see Table 2 for unique R^2 of covariates). Participants did not differ significantly across conditions in task value (mean difference = 2.38; $p = 0.236$), after controlling for pretest task value scores, gender, and Lawson's Test scores.

[INSERT FIGURE 2 ABOUT HERE]

[INSERT TABLE 2 ABOUT HERE]

Tests of the third hypothesis yielded one significant difference: treatment group participants showed higher self-efficacy scores at post-test (mean difference = 4.34; $p < .05$; $d = 0.61$; see Figure 3), after controlling for both pretest self-efficacy scores and lab report performance scores, as well as gender and Lawson's Test scores. Unique variance accounted for by each of the covariates differed somewhat between the treatment and control conditions, with collective performance (i.e., scores for all 4 performance factors) accounting for 0.1% of variance in post-instruction self-efficacy in the treatment condition and 2.4% of variance in the control condition (see Table 3 for unique R^2 of covariates). It demonstrates that participants who received the CTA-based instruction and consequently experienced less extraneous cognitive load, increased their levels of self-efficacy, independent of their performance.

[INSERT FIGURE 3 ABOUT HERE]

[INSERT TABLE 3 ABOUT HERE]

Discussion

Cognitive load theory has consistently treated motivational beliefs solely as a precursor to instruction rather than a possible consequence of instructional design grounded in CLT principles. The purpose of this study was to test the hypothesis that instructional conditions affecting the cognitive load imposed on a learner can directly impact post-instruction motivational beliefs that do not derive from performance. The results generally support this hypothesis. Specifically, the imposition of more extraneous load predicted lower levels of post-instruction self-efficacy, in contrast to the predictions of Schnotz and colleagues (2009).

Established views of the relationship between motivation and the investment of mental effort have maintained that beliefs about necessary effort drive the subsequent investment of effort (e.g., Cennamo, 1993; Clark, 1999; Salomon, 1983, 1984). Recent studies support these views, consistently finding that tasks for which participants had lower self-efficacy were the ones to which they allocated increased effort, leading to stronger task performance (Yeo & Neal, 2008). In instances where participants did not have past performance experiences to draw upon, self-efficacy beliefs positively predicted performance (Sitzmann & Yeo, 2013).

In the current study, participants had not written a full laboratory report for the course prior to the assignment whose scores were analyzed here. Further, it was the first laboratory course in the biology sequence, so it is likely to have been their first such experience at the undergraduate level. Participants did not differ significantly across conditions in their self-efficacy or task value beliefs at the outset of the course, and there was no reason to expect that students would hold differing effort expectations based on the format of the instruction, because they were blind to the existence and nature of the differences between the treatment and the control conditions. Further, performance on the report accounted for minimal variance on post-instruction self-efficacy measures (0.01% in the treatment condition; 2.4% in the control condition), leaving the manipulation of cognitive load through instruction as the most likely source of the medium-large effect size (d = 0.61; Cohen, 1988) effect of the treatment on self-efficacy. Thus, we conclude that the level of extraneous cognitive load experienced during instruction may shape self-efficacy beliefs independent of performance.

In interpreting these findings, it is possible that participants consciously considered mental effort necessary during instruction as a function of cognitive load when reporting their post-instruction self-efficacy. However, it is also possible that participants' assessment of necessary effort and related self-efficacy beliefs engaged unconscious processes (Bargh, Gollwitzer, Lee-Chai, Barndollar, & Trotschel, 2001; Clark, 2014). Extensive research has demonstrated unconscious influences of "cognitive feelings" (Greifeneder, Bless, & Pham, 2011, p. 107), including subjective mental effort (Schwarz & Clore, 2006) and ease of recall (Schwarz, 1998) on subsequent judgments. Typically, judgments preceded by lower effort investment in weighing evidence or greater feelings of ease during recall are associated with more positive evaluations. It is possible that such mechanisms could impact self-efficacy beliefs during or upon reflection of a low extraneous load instructional condition and subsequently impact judgments of self-efficacy more positively than those generated by participants in the high load condition.

Implications and Limitations

The findings reported here have important implications for both the further development of cognitive load theory and practical considerations in instructional design. From a theoretical perspective, the presence of motivational outcomes distinct from learning gains as a function of cognitive load manipulation suggest a need for future research to further verify and explore a greater range of interactions between cognitive load and motivation. The possibility that the cognitive load imposed on working memory during instruction can have direct impacts on motivational beliefs introduces opportunities to connect historically distinct theoretical frameworks. Cognition and motivation may not simply interact through a convergence of independent mechanisms. The processes that give rise to each may be more fundamentally entwined.

There are several specific implications for future research to address additional questions, as well as the limitations of the current study's design. First, additional studies are necessary to replicate the observed effects and, should they be successful, investigate further the possible influence of cognitive load types. In the present study, CLT would characterize the instructional manipulation as imposing additional extraneous load on control participants by delivering task-specific instruction that provided fewer details relevant to attaining stronger performance on the laboratory report task (Author, 2007; Author et al., 2009; Kirschner et al., 2006). However, differentiating between types of cognitive load (i.e., intrinsic, extraneous, germane) can be problematic for both practical and theoretical reasons (de Jong, 2010; Kalyuga, 2011; Schnotz & Kurschner, 2011; Sweller, 2010). As such, more specific interventions must target manipulations of different types of load to determine if the effect is specific to extraneous load or to the overall level of cognitive load imposed by instruction (e.g., Likourezos & Kalyuga, 2017).

Second, although the double-blind quasi-experimental design, the naturalistic environment, and the large sample size are strengths of the current study (Lazowski & Hulleman, 2015), several aspects of the design limit the conclusions that can be drawn. First, this study did not collect cognitive load data from the participants in the course, so it cannot definitively confirm that the treatment did maximize cognitive efficiency (i.e. reduce cognitive load relative to performance-relevant content) for learners in the treatment condition relative to those in the control. Although the anticipated difference in post-instruction performance was attained and the assumption of enhanced cognitive efficiency in the treatment condition was supported by laboratory-based results, there is room for alternative explanations. For example, scientific reasoning ability accounted for substantially more variance in performance scores in the control condition than in the treatment, which is consistent with Bloom's (1984) findings that more specific and procedure-focused instruction was associated with a large drop in the correlation between student aptitude and academic achievement (from $r = 0.60$ to $r = 0.25$). While these findings are not incompatible with cognitive load theory, they do not inherently require cognitive load as a construct.

Another limitation of this study was the lack of multiple measured performance events that could have better informed ongoing research related to the nature of the influence of self-efficacy on performance. Consistent with Bandura's claims (1997; Bandura & Locke, 2003), pre-performance self-efficacy was positively associated with performance outcomes. However, the current findings suggest that the positive association of post-instruction self-efficacy and performance in this case was correlative but not causal, with performance accounting for little variance in self-efficacy gains. The observed correlation is thus likely due to the influence of the training condition as the common variable.

Lastly, the implications of this study for instructional designers suggest reducing extraneous cognitive load imposed in order to promote students' self-efficacy, beyond the known benefits for knowledge acquisition. If instructional design considerations can play a role in shaping students' motivational beliefs, applications of CLT can be brought to bear on a broader scope of training issues that transcend outcomes from individual implementations of instruction and potentially impact the students' pursuit of higher level goal attainment. Motivational beliefs are an important outcome of learning experiences "that individuals hold about their abilities and about the outcome of their efforts, [which] powerfully influence the ways...they will behave" (Pajares, 1996, p. 543) over longer spans of academic and professional endeavors.

References

Author. (2007).

Author (2010).

Author et al. (2008).

Author et al. (2009).

Author et al. (2010).

Author et al. (2013).

Aarts, H., Verplanken, B., & van Knippenberg, A. (1998). Predicting behavior from actions in the past: Repeated decision making or a matter of habit? *Journal of Applied Social Psychology, 28,* 1355-1374.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*, 369-406.

Ayres, P., & Paas, F. (2012). Cognitive load theory: New directions and challenges. *Applied Cognitive Psychology*, *26*, 827-832.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*, 122-147.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Macmillan.

Bandura, A. (2007). Much ado over a faulty conception of perceived self-efficacy grounded in faulty experimentation. *Journal of Social and Clinical Psychology*, *26*, 641-658.

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Multifaceted impact of self-efficacy beliefs on academic functioning. *Child Development, 67,* 1206-1222.

Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K., & Trotschel, R. (2001). The automated will: Non-conscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology, 81,* 1014-1027.

Beck, J. W., & Schmidt, A. M. (2015). Negative relationships between self-efficacy and performance can be adaptive: The mediating role of resource allocation. *Journal of Management*, 0149206314567778.

Britt, T. W. (2005). The effects of identity-relevance and task difficulty on task motivation, stress, and performance. *Motivation and Emotion, 29*, 189-202.

Cennamo, K. S. (1993). Learning from video: Factors influencing learners' preconceptions and invested mental effort. *Educational Technology Research and Development*, *41*, 33-45.

Clark, R. E. (1999). The CaNE (Commitment and Necessary Effort) model of work motivation: A two-stage process of goal commitment and mental effort. In J. Lowyck (Ed.), *Trends in corporate training*. Leuven, Belgium: University of Leuven Press.

Clark, R.E. (2009). How much and what type of guidance is optimal for learning from instruction? In: S. Tobias & T.M. Duffy (Eds.), *Constructivist theory applied to instruction: Success or failure?* New York: Routledge, Taylor and Francis.

Clark, R. E. (2014). Resistance to change unconscious knowledge and the challenge of unlearning. In Berliner, D. C., & Kupermintz, H. (Eds.), *Fostering change in institutions, environments, and people: A festschrift in honor of Gavriel Salomon* (pp. 75-94). New York: Routledge.

Cohen, J. (1988). *Statistical power analysis for the behavioral science*s (2nd ed.). Hillsdale, NJ: Erlbaum.

Cowan, N. (2001). Metatheory of storage capacity limits. *Behavioral and Brain Sciences*, *24*, 154-176.

de Jong, T. (2010). Cognitive load theory, educational research, and instructional design: Some food for thought. *Instructional Science, 38,* 105-134.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*, 211.

Gerjets, P., Scheiter, K., & Cierniak, G. (2009). The scientific value of cognitive load theory: A research agenda based on the structuralist view of theories. *Educational Psychology Review*, *21*, 43-54.

Glynn, S. M., Brickman, P., Armstrong, N., & Taasoobshirazi, G. (2011). Science motivation questionnaire II: Validation with science majors and nonscience majors. *Journal of Research in Science Teaching*, *48*, 1159-1176.

Greifeneder, R., Bless, H., & Pham, M. T. (2011). When do people rely on affective and cognitive feelings in judgment? A review. *Personality and Social Psychology Review, 15,* 107-141.

Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology*, *74*, 657-690.

Kalyuga, S. (2011). Cognitive load theory: How many types of load does it really need?. *Educational Psychology Review*, *23*, 1-19.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, *25*, 351-371.

Kirschner, P., Sweller, J., & Clark, R. E. (2006). Why minimally guided learning does not work: An analysis of the failure of discovery learning, problem-based learning, experiential learning and inquiry-based learning. *Educational Psychologist*, *41*, 75-86.

Knight, J. K., & Smith, M. K. (2010). Different but equal? How nonmajors and majors approach and learn genetics. *CBE-Life Sciences Education*, *9*, 34-44.

Lawson, A. E. (1978). The development and validation of a classroom test of formal reasoning. *Journal of Research in Science Teaching*, *15*, 11-24.

Lawson, A. E. (2000). The generality of hypothetico-deductive reasoning: Making scientific thinking explicit. *The American Biology Teacher*, *62*, 482-495.

Lawson, A. E., Alkhoury, S., Benford, R., Clark, B. R., & Falconer, K. A. (2000). What kinds of scientific concepts exist? Concept construction and intellectual development in college biology. *Journal of Research in Science Teaching, 37,* 996-1018.

Lazowski, R. A., & Hulleman, C. S. (2015). Motivation interventions in education: A meta-analytic review. *Review of Educational Research*, *86*, 602-640.

Lewis, M. W., Bishay, M., McArthur, D., & Chou, J. (1993). Supporting discovery learning in mathematics: design and analysis of an exploration environment and inquiry activities. *Instructional Science*, *21*, 473-88.

Likourezos, V., & Kalyuga, S. (2017). Instruction-first and problem-solving-first approaches: Alternative pathways to learning complex tasks. *Instructional Science*, *45*, 195-219.

Masunaga, H., & Horn, J. (2000). Characterizing mature human intelligence: Expertise development. *Learning and Individual Differences*, *12*, 5-33.

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, *93*, 187-198.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*, 43-52.

Moos, D. C. (2009). Note-taking while learning hypermedia: Cognitive and motivational considerations. *Computers in Human Behavior*, *25*, 1120-1128.

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review*, *19*, 309-326.

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology, 38,* 30-38.

Nie, Y., & Lau, S. (2010). Differential relations of constructivist and didactic instruction to students' cognition, motivation, and achievement. *Learning and Instruction*, *20*, 411-423.

Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, *84*, 429.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, *38*, 1-4.

Paas, F., Tuovinen, J. E., van Merriënboer, J. J., & Darabi, A. A. (2005). A motivational perspective on the relation between mental effort and performance: Optimizing learner involvement in instruction. *Educational Technology Research and Development*, *53*, 25-34.

Paas, F., & van Merriënboer, J. J. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors*, *35*, 737-743.

Pintrich, P. R. (1990). Implications of psychological research on student learning and college teaching for teacher education. Handbook of research on teacher education, 826-857.

Pintrich, P. R., Smith, D., Garcia, T., & McKeachie, W. (1991). The motivational strategies for learning questionaire (MSLQ). Ann Arbor: University of Michigan.

Pintrich, P. R., Smith, D. A., García, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement*, *53*, 801-813.

Rey, G. D., & Buchwald, F. (2011). The expertise reversal effect: cognitive load and motivational explanations. *Journal of Experimental Psychology: Applied*, *17*, 33.

Richard, E. M., Diefendorff, J. M., & Martin, J. H. (2006). Revisitng the within-person self-efficacy and performance relation. *Human Performance, 19,* 67-87.

Rikers, R. M., Schmidt, H. G., & Boshuizen, H. P. (2000). Knowledge encapsulation and the intermediate effect. *Contemporary Educational Psychology*, *25*, 150-166.

Rosenthal, R. (1966). *Experimenter effects in behavioral research*. New York: Appleton-Century-Crofts.

Salomon, G. (1983). The differential investment of mental effort in learning from different sources. *Educational Psychologist, 18,* 42-50.

Salomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology, 76*, 647-658.

Schnotz, W., Fries, S., & Horz, H. (2009). Motivational aspects of cognitive load theory. In M. Wosnita, S. A. Karabenick, A. Efklides, & P. Nenniger (eds.), Contemporary motivation research: From global to local perspectives (pp. 69-96). Cambridge, MA: Hogrefe & Huber Publishers.

Schnotz, W., & Kurschner, C. (2007). A reconsideration of cognitive load theory. *Educational Psychology Review, 19,* 469-508.

Schunk, D. H., Pintrich, P. R., & Meece, J. L. (1996). *Motivation in Education: Theory, Research and Applications* 3rd ed. Englewood Cliffs, NJ: Prentice Hall Merrill.

Schwarz, N. (1998). Accessible content and accessibility experiences: The interplay of declarative and experiential information in judgment. *Personality and Social Psychology Review, 2,* 87-99.

Schwarz, N., & Clore, G. L. (2006). Feelings and phenomenal experiences. In E. T. Higgins & A. Kruglanski (Eds.), *Social psychology. A handbook of basic principles (2nd ed.).* New York: Guilford Press.

Sitzmann, T., & Yeo, G. (2013). A meta-analytic investigation of the within-person self-efficacy domain: Is self-efficacy a product of past performance or a driver of future performance? *Personnel Psychology*, *66*(3), 531-568.

Spanjers, I. A. E., van Gog, T., & van Merriënboer, J. J. G. (2012). Segmentation of worked examples: Effects on cognitive load and learning. *Applied Cognitive Psychology*, *26*, 352-358.

Sullivan, M. E., Yates, K. A., Inaba, K., Lam, L., & Clark, R. E. (2014). The use of cognitive task analysis to reveal the instructional limitations of experts in the teaching of procedural skills. *Academic Medicine*, *89*, 811-816.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*, 257-285.

Sweller, J. (1993). Some cognitive processes and their consequences for the organization and presentation of information. *Australian Journal of Psychology, 45*, 1-8.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*, 295-312.

Sweller, J. (2004). Instructional design consequences of an analogy between evolution by natural selection and human cognitive architecture. *Instructional Science*, *32*, 9-31.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review, 22,* 123-138.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*, 251-296.

Timmerman, B. E. C., Strickland, D. C., Johnson, R. L., & Payne, J. R. (2011). Development of a 'universal's rubric for assessing undergraduates' scientific reasoning skills using scientific writing. *Assessment & Evaluation in Higher Education*, *36*, 509-547.

Tobias, S. (1994). Interest, prior knowledge, and learning. *Review of Educational Research*, *64*, 37-54.

Tuovinen, J. E., & Paas, F. (2004). Exploring multidimensional approaches to the efficiency of instructional conditions. *Instructional Science*, *32*, 133-152.

van Gog, T., & Rummel, N. (2010). Example-based learning: Integrating cognitive and social-cognitive research perspectives. *Educational Psychology Review, 22,* 155-174.

van Merriënboer, J. J., & Ayres, P. (2005). Research on cognitive load theory and its design implications for e-learning. *Educational Technology Research and Development*, *53*, 5-13.

van Merriënboer, J. J., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, *17*, 147-177.

Vicente, K. J., & Wang, J. H. (1998). An ecological theory of expertise effects in memory recall. *Psychological Review*, *105*, 33-57.

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary educational psychology*, *25*, 68-81.

Yeh, Y. Y., & Wickens, C. D. (1988). Dissociation of performance and subjective measures of workload. *Human Factors*, *30*, 111-120.

Yeo, G., & Neal, A. (2008). Subjective cognitive effort: A model of states, traits, and time. *Journal of Applied Psychology*, *93*, 617-631.

Zheng, R., McAlack, M., Wilmes, B., Kohler-Evans, P., & Williamson, J. (2009). Effects of multimedia on cognitive load, self-efficacy, and multiple rule-based problem solving. *British Journal of Educational Technology, 40,* 790-803.

Figure 1. Adjusted mean scores for the Discussion factor in lab reports after controlling for gender and Lawson's Test scores. The mean difference is significant at 0.45, $p < .05$ ($d = 0.36$).

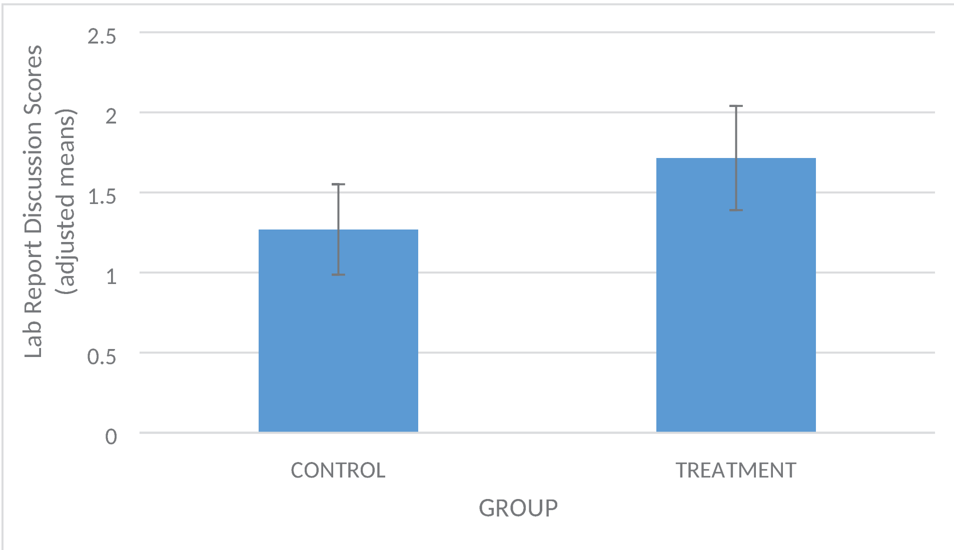

Figure 2. Adjusted means for post-instruction self-efficacy scores after controlling for pretest self-efficacy scores, gender, and Lawson's Test scores. The mean difference is nonsignificant using a 2-tailed test at 3.90, p = 0.093 (*d* = 0.54).

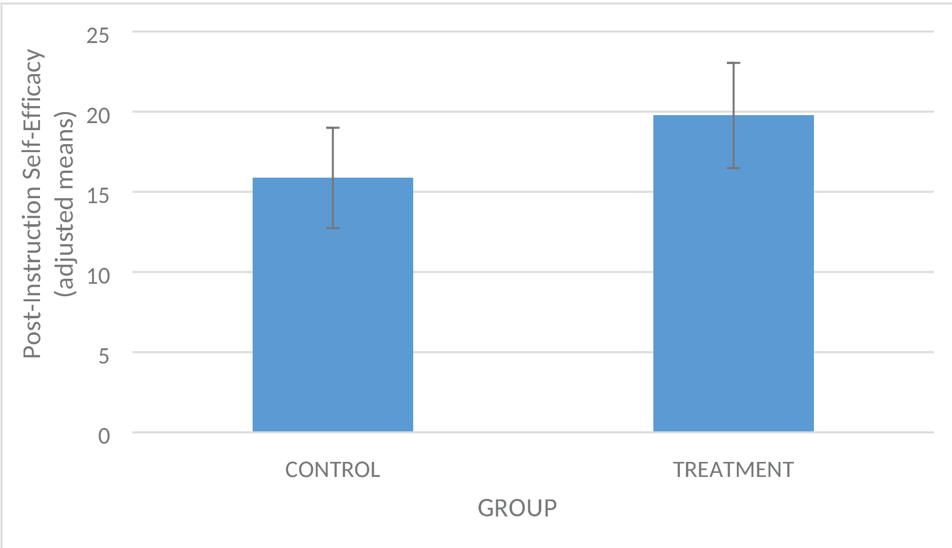

Figure 3. Adjusted mean post-instruction self-efficacy scores after controlling for both pretest self-efficacy scores and lab report performance scores, as well as gender and Lawson's Test scores. The mean difference is significant at 4.34, $p < .05$ ($d = 0.61$).

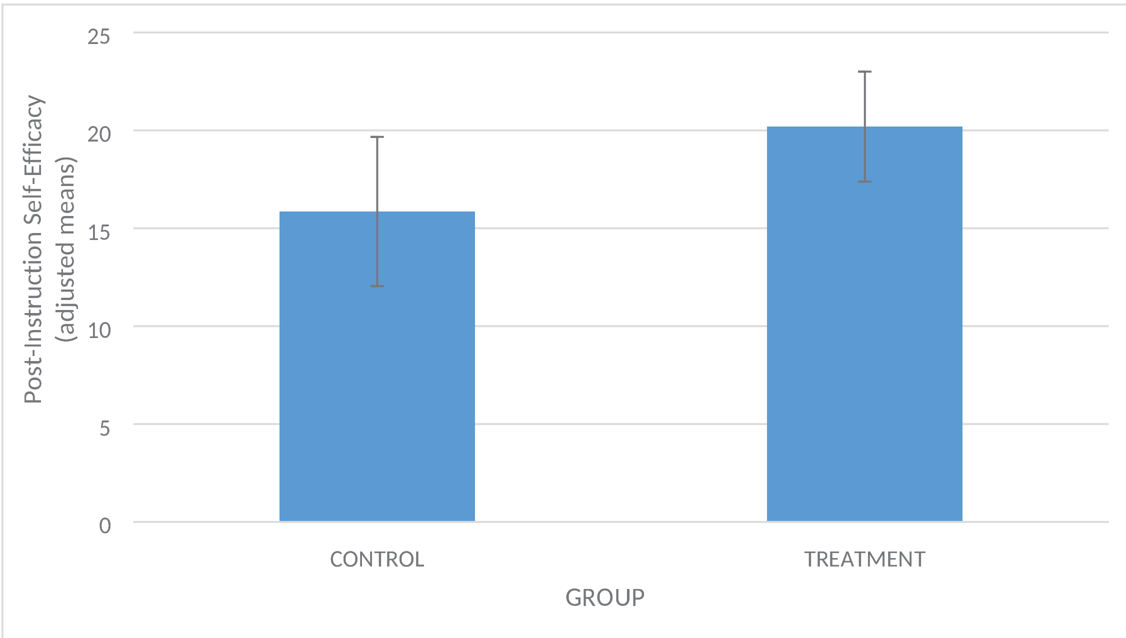

Table 1.

Dependent Variable	Covariates	Treatment		Control	
		Unique Variance	Cumulative Variance	Unique Variance	Cumulative Variance
Framing the Study factor score	Gender	0.006	0.006	0.01	0.01
	Lawson's Test score	0.006	0.012	0.012	0.022
Hypotheses factor score	Gender	0.001	0.001	0.001	0.001
	Lawson's Test score	0.011	0.012	0.014	0.015
Results factor score	Gender	0.000	0.000	0.000	0.000
	Lawson's Test score	0.017	0.017	0.039	0.039
Discussion factor score	Gender	0.001	0.001	0.000	0.000
	Lawson's Test score	0.011	0.012	0.043	0.043

Table 2.

Dependent Variables	Covariates	Treatment		Control	
		Unique Variance	Cumulative Variance	Unique Variance	Cumulative Variance
Post-Instruction Self-Efficacy score	Pre-Instruction Self-Efficacy score	0.097	0.097	0.090	0.090
	Gender	0.093	0.190	0.087	0.177
	Lawson's Test score	0.093	0.283	0.100	0.277
Post-Instruction Task Value score	Pre-Instruction Task Value score	0.128	0.128	0.120	0.120
	Gender	0.130	0.258	0.128	0.248
	Lawson's Test score	0.130	0.388	0.128	0.376

Table 3.

Covariates	Treatment		Control	
	Unique Variance	Cumulative Variance	Unique Variance	Cumulative Variance
Pre-instruction Self-Efficacy score	0.073	0.073	0.080	0.080
Gender	0.001	0.073	0	0.082
Lawson's Test score	0.001	0.073	0.016	0.095
Framing the Study factor score	0	0.073	0.018	0.105
Hypotheses factor score	0	0.073	0.002	0.112
Results factor score	0	0.073	0.002	0.112
Discussion factor score	0.001	0.074	0.002	0.113