View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by ASU Digital Repository

Exploring the Use of Self-explanation Prompts in a Collaborative Learning Environment

by

Kyle Matthew Wright

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2018 by the
Graduate Supervisory Committee:

Robert K. Atkinson, Chair
Wilhelmina C. Savenye
Brian Nelson

ARIZONA STATE UNIVERSITY

May 2018

ABSTRACT

A recorded tutorial dialogue can produce positive learning gains, when observed and used to promote discussion between a pair of learners; however, this same effect does not typically occur when an leaner observes a tutorial dialogue by himself or herself. One potential approach to enhancing learning in the latter situation is by incorporating self-explanation prompts, a proven technique for encouraging students to engage in active learning and attend to the material in a meaningful way. This study examined whether learning from observing recorded tutorial dialogues could be made more effective by adding self-explanation prompts in computer-based learning environment. The research questions in this two-experiment study were (a) *Do self-explanation prompts help support student learning while watching a recorded dialogue?* and (b) *Does collaboratively observing (in dyads) a tutorial dialogue with self-explanation prompts help support student learning while watching a recorded dialogue?* In Experiment 1, 66 participants were randomly assigned as *individuals* to a physics lesson (a) with self-explanation prompts (Condition 1) or (b) without self-explanation prompts (Condition 2). In Experiment 2, 20 participants were randomly assigned in 10 pairs to the same physics lesson (a) with self-explanation prompts (Condition 1) or (b) without self-explanation prompts (Condition 2). Pretests and posttests were administered, as well as other surveys that measured motivation and system usability. Although supplemental analyses showed some significant differences among individual scale items or factors, neither primary results for Experiment 1 or Experiment 2 were significant for changes in posttest scores from pretest scores for learning, motivation, or system usability assessments.

ACKNOWLEDGMENTS

This work is dedicated to Beverly Wright, David Wright, and Danielle Wright. Without the motivation and love of my mother, father, and sister none of this would have been possible.

Along the journey, there have been many supporters to thank. Chief among those who have supported me at every step of the way has been my advisor and committee chair Dr. Robert Atkinson. Dr. Wilhelmina Savenye and Dr. Brian Nelson, I appreciate your input, guidance, and encouragement.

TABLE OF CONTENTS

Page

LIST OF TABLES ... vii

LIST OF FIGURES ... ix

CHAPTER

1 INTRODUCTION ... 3

Purpose of This Study ... 3

Research Question ... 3

Literature Review ... 4

Self-explanation Prompts ... 4

Interactive, Constructive, Active, and Passive Framework ... 11

Measuring System Usability and Intrinsic Motivation in a Learning Environment ... 16

Cognitive Load Theory ... 17

Overview of This Study ... 17

2 METHOD ... 19

Experiment 1 ... 19

Participants and Design ... 19

Learning Environment and Materials ... 20

Measures ... 24

Procedure ... 26

Scoring ... 27

Experiment 2 ... 28

CHAPTER Page

Participants and Design 28

Learning Environment 29

Measures 29

Procedure 30

Scoring 30

3 RESULTS 31

Experiment 1 31

Research Question 1 31

System Usability Scale Primary Analysis Results 32

Intrinsic Motivation Inventory Primary Analysis Results 32

Experiment 1 Supplemental Analyses 33

Paired-samples *t*-test to Evaluate Learning Gains in Experiment 1 33

System Usability Scale Supplemental Analysis Results 34

Intrinsic Motivation Inventory Supplemental Analysis Results 35

Interpretation of Supplemental Analysis for Experiment 1 36

Experiment 2 36

Research Question 2 37

System Usability Scale Primary Analysis Results 37

CHAPTER Page

Intrinsic Motivation Inventory Primary Analysis Results .. 38

Experiment 2 Supplemental Analyses 38

Paired-samples *t*-test to Evaluate Learning Gains in Experiment 2 38

System Usability Scale Supplemental Analysis Results..... 39

Intrinsic Motivation Inventory Supplemental Analysis Results 40

Interpretation of Supplemental Analysis for Experiment 2 41

4 CONCLUSIONS AND DISCUSSION 42

Experiment 1 42

Experiment 2 44

Limitations 44

Future Research 45

REFERENCES 48

APPENDIX

A SYSTEM USABILITY SCALE 52

B INTRINSIC MOTIVATION INVENTORY 54

C ITEMS PRESENTED IN THE TUTORING SESSION 57

D PRETEST ITEMS 59

E POSTTEST ITEMS 61

F DEMOGRAPHIC SURVEY TIEMS 63

APPENDIX Page

G SUMMARY OF THE EXPERIMENTS 65

LIST OF TABLES

Table Page

1. Summary of Participants according to their Self-reported Majors in Experiment 1 20
2. Justification and Self-focused Self-explanation Prompts Used in the Learning Environment 24
3. Possible Scores for Intrinsic Motivation Inventory Subscales 28
4. Summary of Participants according to their Self-reported Majors in Experiment 2 29
5. Learning Pretest and Posttest Means and Standard Deviations (and Posttest Means and Standard Deviations by Condition) for Experiment 1 31
6. Independent *t*-test Results (Equal Variances Assumed) for Experiment 1 32
7. The Means and Standard Deviations by Condition for System Usability Scale Scores in Experiment 1 32
8. The Means and Standard Deviations by Condition for Intrinsic Motivation Inventory Scores in Experiment 1 33
9. Paired-samples *t*-test Results (Equal Variances Assumed) for Experiment 1 34
10. Independent *t*-test Results (Equal Variances Assumed) for System Usability Scale Items for Experiment 1 35
11. Independent-samples *t*-test Results (Equal Variances Assumed) for Intrinsic Motivation Inventory Factors for Experiment 1 36
12. Learning Pretest and Posttest Means and Standard Deviations (and Posttest Means and Standard Deviations by Condition) for Experiment 2 37

Table Page

13. Independent *t*-test Results (Equal Variances Assumed) for Average System Usability Scale Scores for Experiment 2 37

14. Independent *t*-test Results (Equal Variances Assumed) for Average Intrinsic Motivation Inventory Scores for Experiment 2 38

15. Paired-samples *t*-test Results (Equal Variances Assumed) for Experiment 2 39

16. Independent *t*-test Results (Equal Variances Assumed) for System Usability Scale Items for Experiment 2 40

17. Non-significant Independent *t*-test Supplemental Analysis Results (Equal Variances Assumed) for Intrinsic Motivation Inventory Factors for Experiment 2 41

LIST OF FIGURES

Figure Page

1. Screenshot of Tutoring Segment in Learning Environment 21

2. Self-explanation Prompt in the Learning Environment 22

CHAPTER 1

Introduction

The ubiquity of technology, digital tools, and digital media has expanded in the last few decades. Computer-based instruction has become more available to learners at an almost exponential growth rate. Digital media can be reused or recycled in a computer-based learning environment; this practical use of digital media has had an impact on the research into educational technology and the development of educational technology tools (Muldner, Lam, & Chi, 2013). One such tool, that is both practical and impactful, is digital video. The incorporation of digital video into computer-based instruction has led to the creation of learning environments that are scalable and effective (Ronchetti, 2010). In particular, there is modest research evidence that document the positive effects of pairs of learners collaboratively observing a tutoring session (Chi, 2009; Chi, Roy & Hausman, 2008; Muldner et al., 2013). Chi et al. (2008) recorded video and audio in all conditions.

In their seminal study on self-explanations, Chi, Bassok, Lewis, Reimann, and Glaser (1989) observed that successful students generated twice the amount of explanations when reviewing example solutions. To determine the impact of the reviewed sample solutions, explanations were defined as statements other than the first reading of the example line or any conversation that does not pertain to the subject matter. Furthermore, the ideas generated from the example solutions can be classified as (a) *explanations*, (b) *monitoring statements*, (c) *other*. Chi et al. (1989) defined explanation statements as "inferences about the conditions, the consequences, the goals, and the meaning of various mathematical actions described in the example" (p. 24). Monitoring statements can be described as any statements in which learners acknowledges their understanding of a subject matter, whereas other statements consisted of learners paraphrasing information, elaborating mathematical operations, or statements relating to student actions. The type of explanation they generated varied. For instance, when viewing a worked-out physics example, Chi et al. found that successful students will sometimes self-monitor their learning by judging their comprehension of the material. Not all students spontaneously produce self-explanation prompts on their own, yet they can be easily created, delivered, and their effects observed through educational technology tools. Moreover, self-explanations can be structured in ways to encourage learners to focus on: (a) the structure or content, (b) comprehension of material, and (c) other attributes (refer to Figure 5 in Chi et al., 1989). Furthermore, the sequential cuing and highlighting of information, within the learning environment, can encourage learners to generate more self-explanations than environments lacking such cues and highlights (De Koning, Tabbers, Rikers, & Paas, 2011). De Koning et al. (2011) found that through animating and highlighting self-explanation prompts, learners gained a deeper understanding of the learning material when compared to their peers learning from a static self-explanation prompt environment. Although the self-explanation literature has numerous studies documenting the use of self-explanation prompts when learning by themselves, their effectiveness in computer-supported collaborative learning environments has not been explored.

One framework that incorporates self-explanation activities and encourages its use is the *interactive/constructive/active/passive* (ICAP) framework (Chi, 2008). Chi (2008) identified four types of activities within the framework as: (a) *passive*, (b) *engaging activities*, (c) *self-construction activities*, and (d) *guided-construction activities* *in instructional dialogue with a peer*. Passive activities are defined as activities in which the learner is not required to perform a type of action, such as not taking notes or reviewing examples while observing a tutorial dialogue. Engaging activities are comprised of a set of active activities such as manipulating videotapes or summarizing dialogue, whereas the purpose of self-construction activities is to elicit new knowledge; examples of self-construction activities include the creation of concept maps or analyzing case studies; however, guided-construction activities in instructional dialogue with a peer is highly interactive; not only can knowledge be constructed through scaffolding the instructional material, but the presence of a peer creates opportunities for knowledge construction. The *ICAP* framework is useful in helping (a) to define and classify engaging learning activities that can be used in a variety of different learning environments and (b) to investigate each mode of engagement and determine whether there are any consistencies within the literature.

Purpose of This Study

This two-experiment study sought to examine whether the use of self-explanation prompts could significantly increase learning for either individuals or learners working collaboratively while observing a tutorial dialogue.

Research Questions

1. Do self-explanation prompts help support students' learning while watching a recorded tutorial dialogue?
2. What effect does collaboratively observing a tutorial dialogue with self-explanation prompts have on learning?

Literature Review

The activity of collaboratively observing dialogue between a tutor and a tutee is not completely interactive; there are moments when dialogue between observers is meaningful; however, most dialogue between observers could be considered constructive (Chi, Leeuw, Chiu & Lavancher, 1994). Chi (2008) defined constructive as a component of the *ICAP* framework, which occurs when learners construct new knowledge based on the given content. Through prompting learners, it is possible to elicit self-explanations, which can result in learning (Chi et al., 1994).

Self-explanation prompts. Chi et al. (1989) explored how individuals construct self-explanations when introduced to new examples in the context of learning by investigating how learners can form generalizations based on new examples. Self-explanations can be described as the process of generating inferences that occur when viewing examples. In the Chi et al. (1989) study, any talk-out-loud statement that was not made in the first reading of the example line was recorded as a self-explanation. Successful learners generated 142 self-explanation statements and spent 13 minutes studying each example; unsuccessful learners generated 21 self-explanations statements and spent 7.4 minutes studying each example; however, when problem-solving, successful learners generated 141 self-explanation statements and spent 13.8 minutes on each problem. Unsuccessful learners generated 122 self-explanation statements and spent 14.3 minutes on each problem.

In a subsequent study, Chi and VanLehn (1991) identified that self-explanations come primarily from two sources. Specifically, self-explanations arise when learners (a) contemplate on the previous text that is initiated by a general statement or procedure in the current example or (b) use example statements to help produce new general knowledge to support their understanding. Chi et al. (1994) further investigated the self-explanation effect and whether it can be achieved through prompting the learner. Moreover, they studied whether learner ability has an impact on the self-explanation effect's ability to produce learning gains. In doing so, this helped to confirm the self-explanation effect by replicating the results of Chi et al. (1989); however, the following changes were made: (a) different domain, (b) different age group of participants, (c) used text rather than worked-out examples, (d) focused on an analytical understanding of concepts, and (e) used prompts to induce self-explanations. The purpose of these changes was to measure the generalization of the self-explanation effect. Chi et al. (1994) produced similar results to the previous study; however, prompting the learner produced 10% higher gains than not prompting the learner.

The concept of self-explanations has been investigated in various learning environments, including in conjunction with example-based learning. Renkl (1997) measured the quality of self-explanation prompts when learning from worked-out examples. Learners who referred to a convention or principle of the subject matter were typically more successful. These two types of self-explanation styles were termed *anticipative reasoning* and *principle-based explanations*. The concept of anticipative reasoning was defined as the calculation of a problem without viewing the example, whereas principle-based explanations were defined as how many times learners referred to a principle. Renkl (1997) posited that with anticipative reasoning learners' self-diagnosis their level of competence. Additionally, the positive learning gains produced through principle-based self-explanation was similar to the findings of Chi et al. (1989).

Berthold, Eysink, and Renkl (2009) discovered that self-explanation prompts that provided learning assistance, such as supporting knowledge, have a positive effect on *procedural knowledge* and *conceptual knowledge*. This is referred to this as the *assisting self-explanation prompt effect*.

Renkl, Stark, Gruber, and Mandl (1998) observed that learners with low prior knowledge benefited from the use of self-explanation prompts. Learners were first presented an example model of self-explaining when observing a worked-out example; additionally learners were coached on the process of self-explaining. Within the study, multiple examples in different contexts or multiple examples in one context were presented, while learners were either prompted to self-explain or not prompted to self-explain. Additionally, examples can be scaffolded and augmented with a variety of cognitive processes. Further research empirically demonstrated an increase in far transfer learning when adding self-explanation prompts to a learning environment, in which an example solution was scaffolded in sequential order from backwards to forwards; this is also known a *backwards fading*. The significant finding on far-transfer revealed that the use of self-explanation prompts led to learners solving problems in the same content area, which were structured and described in a different context with different values (Atkinson, Renkl, & Merrill, 2003). Atkinson et al. (2003) posited that the incorporation of self-explanation prompts would promote active processing, and backwards fading of examples would foster greater anticipative reasoning styles of self-explanation (Renkl, 1997). Atkinson et al. (2003) noted that self-explanation prompts contributed to the success of learners, and with the use of backwards fading, generated a greater number of anticipative reasoning self-explanations. Furthermore, when combining the use of backwards fading worked-examples and self-explanation prompts the combination can significantly increase the quality of learning without increasing the amount of instructional time.

Crippen and Earl (2007) compared the performance between providing worked-examples and worked-examples with self-explanation prompts to the performance of a control group in which the worked-examples and self-explanation prompts were omitted. Comparisons between the conditions indicated worked-examples with self-explanation prompts had a positive effect on performance. Moreover, the combination of worked-examples with self-explanation prompts led to greater *self-efficacy* rather than worked examples alone. Schworm and Renkl (2007) indicated that principle-based prompts can foster self-explanations in learning when students are learning in an unstructured domains or complex skills.

Hilbert and Renkl (2009) compared learning how to create a concept map through practicing and learning through training with *heuristic examples.* A heuristic example was comparable to worked-out examples, because it included a robust example and explanation of a concept map. Through a questionnaire, developed by NASA, that assessed the overall perceived workload of learners the cognitive load imposed on learners was measured. Although self-explanation prompts can contribute to positive learning outcomes, there is also evidence that they can produce a higher germane cognitive load and decrease the extraneous cognitive load. Moreover, there are a variety of different types of self-explanation prompts that have been shown to provide positive effects.

Yeh, Chen, and Hwang (2010) investigated the use of reasoning-based prompts and predicting-based prompts to help generate self-explanations in both lower and higher knowledge learners, and provided empirical evidence for the use of adaptive self-explanation prompts in learning. *Reasoning-based prompts* were defined, according to Renkl (1997), as prompts designed to elicit and define principles which support the actions in the learning environment. *Predicting-based prompts* required learners to anticipate upcoming actions within the animation Comparisons between the reasoning-based and predicting-based prompts took into consideration the level of prior knowledge by separating learners into either high-knowledge or low-knowledge learners and considering scores from a prerequisite test,. Consequently, in this study higher-knowledge learners performed better using predicting-based prompts than reasoning-based prompts, whereas lower-knowledge learners performed better using reasoning-based prompts that predicting-based prompts. Furthermore, the authors also found a significant interaction between the use of prompts and learners' prior knowledge. They found that the level of cognitive load imposed by learning on self-explanation prompts depends on the level of prior knowledge. Additionally, reasoning-based prompts benefited lower-knowledge learners more than higher-knowledge learners albeit higher-knowledge learners benefited the most from reasoning-based prompts (Yeh et al., 2010).

Nokes, Hausmann, VanLehn, and Gershman (2011) determined that *gap-filling prompts*, such as *justification prompts* or *step-focused prompts*, led to generating more self-explanations than did *mental-model revision prompts*. The purpose of gap-filling prompts was to fill in missing information within the example. In instances of justification prompts, learners were asked to identify and focus on the background concepts in each step to take. Mental-model revision prompts were designed to elicit the learners' prior knowledge. Within the current study, an example of a justification prompt was "what principle applies here", whereas an example of a step-focused prompt used in the current study was "What is the direction of the Normal Force acting on the block? What is the value of the Normal Force when (a) the string is taut, (b) the string is cut loose and the box is sliding down." Gap-filling prompts can lead to better performance in problem-solving environments: Justification prompts elicit principles that validate problem-solving steps, whereas step-focused prompts direct attention to explain the details and actions taken within each step. Their findings revealed that learners using justification or step-focused prompts outperformed learners using direct instructional prompts (Nokes et al., 2011).

Different types of prompts that can be used in learning and instruction to augment the generation of self-explanations have been explored in different computer-based learning environments. Lin and Atkinson (2013) observed that the combination of prompts and animated visuals contributed to positive learning outcomes. Furthermore, the authors concluded that "although technologies advance in a surprising speed, instructional designers should pay attention to learners' cognitive aspects, and utilize the combination of technology-based visualizations and cognitive strategies (such as self-explanation prompts) in the design and development of the learning environment" (Lin & Atkinson, 2013, p. 107). In another study, O'Neil et al. (2014) determined that although self-explanation prompts can lead to positive learning outcomes, they can also be a hindrance to learning in computer-based games. The authors noted the most effective types of self-explanation prompts as being those that helped to make connections between terminologies. Conversely, the ineffective self-explanation prompts addressed elementary or esoteric questions or elements. Adams and Clark (2014) found that for self-explanation prompts to be effective cognitive load must be successfully managed; however, self-explanation prompts have been shown to have a positive effect on learning when paired with multiple graphical representations (Rau, Aleven, & Rummel, 2015).

Subsequent research into self-explanation prompts and computer-based learning environments incorporated self-explanation prompts into collaborative environments and digital games. Hsu, Tsai, and Wang (2016) compared the effects of using self-explanation prompts in a single-user gaming environment to using self-explanation prompts in a collaborative gaming environment. The author's determined that self-explanation prompts alone did not adequately impact learning gains; however, self-explanation prompts can foster engagement, which can lead to learning. The ability for self-explanation prompts to adapt and adjust according to learner performance can result in positive learning gains in complex digital environments. Self-explanation prompts in a learning environment, which increased from simple navigational instructions to navigational instructions in terms of Newtonian Mechanics, produced better learning on a posttest than self-explanation prompts that are solely navigational instructions in terms of Newtonian Mechanics (Clark, Virk, Barnes, & Adams, 2016).

Renkl et al. (1998) presented examples in multiple contexts, in which learners negligibly performed better on far-transfer problems. Additionally, learners without the proper guidance to foster self-explanations, failed to achieve high learning outcomes. To overcome this drawback, when eliciting self-explanations, examples within one context should not overburden the learner. In a subsequent study, Renkl (2002) proposed a set of principles for incorporating instructional explanations into learning through self-explanations. He suggested that there are components of the *Self-Explanation Activity Supplemented by Instructional Explanations* principles that must be further adopted when complementing instructional explanations with self-explanations: A high frequency of instructional explanations and self-explanations must be incorporated into the instruction and considerations must be managed to reduce extrinsic cognitive load. Schworm and Renkl (2006) examined the impact of instructional explanations on learners' ability to self-explain in a computer-based learning environment; the instructional explanations were responses to the self-explanation prompt. The addition of instructional explanations resulted in the learner producing less written self-explanation responses. Consequently, the use of instructional prompts had a detrimental effect to generating self-explanations.

Interactive, constructive, active, and passive framework in a tutoring environment. Chi (2009) defined the terms *active*, *constructive, interactive*, and noted that although these terms are used frequently in the literature, they are seldom defined. Activities, used in studies within the literature, can be classified into *engaging activities, self-construction activities* or *guided-construction activities in instructional dialogue*. Active can be defined as learners engaging in action during an instructional activity, whereas constructive can be described as building new knowledge from given content, including activities such as speaking with another person or interacting with a learning environment. Through analyzing the literature and characterizing the terms active, constructive, and interactive, a testable hypothesis was developed "that overall, active is better than passive, constructive is better than active, and interactive is better than constructive" (Chi, 2009, p. 88). The three ways of testing this hypothesis were (a) to break down all comparisons into pairs, (b) to determine whether identical activities produced similar learning outcomes, and (c) to use a specific activity and contrast each condition.

Chi et al. (2008) compared human tutoring to the following conditions: (a) observing collaboratively, (b) collaborating, (c) observing alone, and (d) studying alone. The authors found support for human tutoring. Additionally observing collaboratively and observing alone was just as effective as human tutoring. Interactions between observers of a tutorial dialogue can have a positive impact on learning; however, there are difficulties in engaging in constructive learning that could arise from the passiveness of observations. The results revealed support for the active/constructive/interactive observing hypothesis in the following ways: (a) collaboratively interacting with an observer was just as effective as participating in a tutoring session; (b) learning increased as collaborative observers interacted with one another; (c) collaboratively observing produced higher learning gains than the lone observing condition; and (d) active lone observers who manipulated the tutoring tape (e.g. rewind, pause, or faster forward) and posed questions aloud, produced higher learning gains than passive loan observers who did not engage in manipulating the tutoring tape nor posed questions aloud. This finding further supported the active/constructive/interactive observing hypothesis.

The active, constructive, and interactive hypothesis was investigated by Muldner et al., (2014). Specifically the authors investigated and compared: (a) *collaboratively observing dialogue*, (b) *one-on-one tutoring*, and (c) *collaboratively observing monologue* at a southwestern university using undergraduates. Subsequently, identical conditions were compared using a younger population in a second study. Both studies incorporated the science topic of molecular diffusion; the authors noted the misconceptions and challenges that existed within understanding this topic. Within each study, learning gains were calculated and substantive contributions were analyzed. Substantive contributions were determined by using the convention set up in Chi et al. (2008). Learners in the collaboratively observing dialogue condition did not significantly perform better on the posttest than learners in the one-on-one tutoring condition; however, learners in the collaboratively observing dialogue condition produced more meaningful interactions.

In both studies, learners in the one-on-one tutoring condition generated more substantive contributions than those in the observing conditions. Learners in the collaboratively observing dialogue condition significantly produced more substantive contributions than those in the monologue observer condition; however, in the second study, this was not significant. Furthermore, the significant substantive contributions of learners in the one-on-one tutoring condition did not produce significant learning gains in the second study.

Chi, Kang, and Yaghomourian (2017) observed and investigated how content movement of a tutor or tutee influences learning when observing a tutorial dialogue and observing a tutorial monologue. Content moves were defined as information that is said by either the tutor or tutee. The content movement of subjects in the tutorial dialogue videos and in the tutorial monologue videos was determined; tutorial dialogues contained unique content moves. Within the tutorial dialogue videos there were five unique tutor and tutee content moves and three common content moves. The unique tutor and tutee content moves were the following: (a) tutor gave elaborate feedback to the tutee, (b) tutor asked a deep question, (c) tutor gave an incorrect statement, (c) substantive comments of the tutee, and (d) questions raised by the tutee. Common content moves were iconic gestures, deictic gestures, and tutor covered concepts. None of the unique or common moves was significantly correlated with learning outcomes.

Chi et al. (2017) further analyzed the interactions between the tutorial dialogue video observers, tutorial monologue video observers, and participants' workbooks to investigate how reliable the ICAP hypothesis was at predicting how well students learn when they are active. The comments of the observing dyads in the tutorial dialogue and the tutorial monologue conditions were defined in terms of these: (a) interactive, (b) constructive, and (c) active. To be considered a constructive comment, the observer had to refer to a topic or subject being covered in the tutoring videos, whereas interactivity was composed of more than one constructive comment on the same topic or subject. Regarding the participant's workbooks, there were a total of 22 items to accompany the tutorial dialogue and tutorial monologue videos. The problems presented in the workbook could either be solved or copied from the tutorial videos; problems that were solved reflected a constructive activity, whereas problems that were copied reflected an active activity. Tutorial dialogue video observers significantly solved more problems than did tutorial monologue video observers. This finding is supported by the ICAP hypothesis that a constructive behavior produces more learning than does an active behavior. Moreover, constructive comments were significantly correlated with learning gains.

Menske, Stump, Krause, and Chi (2013) applied the *Differentiated Overt Learning Activities framework* (DOLA) to the learning domain of engineering. The purpose of DOLA is to divide overt active learning methods into interactive, constructive, or active activities. The interactive activities included (a) completing a partially drawn concept map, (b) matching historical events with scientific reasons for their occurrence, and (c) analyzing the properties of a unit cell, whereas the constructive activities were (a) create and calculate the indices of a unit cell planes and (b) create the atoms in a unit cell and determine how many atoms are in three-unit cells. The active activities compared were (a) selection of materials, (b) copying unit cell directions, and (c) copying the unit cell directions for families. After completing each activity, student learning was measured using items that relied on recall of information (*verbatim*), integration of multiple ideas and information (*integration*), and the construction of new concepts or ideas (*inference*). Even though there were no significant differences between interactive and constructive activities, there was a significant positive difference in the inference scores for interactive activities. This finding also provided evidence to support the ICAP hypothesis developed by Chi (2009).

Based on the results from the first experiment, Menske, Stump, Krause, and Chi (2013) investigated the differential effects among four conditions that corresponded to the DOLA framework and the ICAP hypothesis in a controlled laboratory setting in a second experiment. The four conditions were (a) interactive, (b) constructive, (c) active, and (d) passive. A pairwise comparison determined that, overall, learners in the interactive condition outperformed learners in all other conditions. Additionally, learners in the constructive condition outperformed those in the passive condition, and those in the active condition outperformed those in the passive condition. Consequently, there is support for utilizing interactive learning activities.

Chi and Wylie (2014) further described the active, constructive, and interactive hypothesis and defined the ICAP framework. The ICAP framework proposed that behaviors can be organized and categorized into active, constructive, interactive, or passive. Furthermore, as engagement increases, from passive, to active, to constructive, and to interactive, different cognitive processes occurred that increased learning. Although passive engagement requires storing new knowledge it does not require that the new knowledge be integrated with prior knowledge. Conversely, through active engagement individuals manipulate information, thereby activating prior knowledge. Chi and Wylie (2014) described the impact of active engagement as “quite substantial because significant knowledge completion has occurred. One could say that they have achieved, at a minimum, a shallow understanding” (p. 10). Constructive engagement led to the construction and linking of schemata, which can dramatically alter knowledge structure. Through interactive engagement, new schemata are created and linked in a continuous pattern, which are further strengthened by contributions from a peer.

Measuring system usability and intrinsic motivation in a learning environment. Brooke (1996) originally designed the *System Usability Scale* (SUS) to measure the usability of a variety of different products, including websites and applications; the scale consists of 10 subjective statements to which subjects responded to on a Likert Scale. Half of the statements are positive, and half of the statements are negative statements. The scores is calculated by subtracting 1 from the scores for items 1, 3, 5, 7, and 9, and subtracting five from the score for items 2, 4, 6, 8, and 10. Last, this number is multiplied by 2.5. The SUS can be viewed in Appendix E.

Ryan and Deci (2000) created The *Intrinsic Motivation Inventory* (IMI) that consists of 45 subjective statements and measured the following on a Likert scale: (a) interest and enjoyment, (b) perceived competence, (c) effort, (d) value, (e) pressure and tension, and (f) choice while performing activity. The scores are calculated by reverse scoring 16 of 45 items. The possible scores for the IMI subscale are in Appendix F.

Cognitive Load Theory. Cognitive Load Theory proposes that knowledge stored in working memory or long-term memory can be categorized as either intrinsic cognitive load, extraneous cognitive load, or germane cognitive load. Intrinsic cognitive load was a measure of the elements pertaining to the content or knowledge (Pass, Renkl, & Sweller, 2003). Sweller (1994) posited that low intrinsic cognitive load would have a low number of elements interacting. Extraneous cognitive load can be defined as information that is unnecessary and can interfere with knowledge acquisition. Germane cognitive load was a measure of how the information is presented to the learner (Pass et al., 2003). Pass et al. (2003) asserted that instructional designers can highly influence germane cognitive load.

Overview of This Study

The initial design of this experiment was a 2x2 ANOVA between-subjects experiment with two independent variables, each with two levels. The first independent variable was presence of self-explanation prompts (*present* or *absent*), and the second variable was the type of learning (*individually* or *collaboratively through dyads*). It was designed to determine whether there was (a) a main effect for self-explanation prompts on learning, (b) a main effect for collaborative learning on learning, or (c) an interaction between the two. Complications arose around the complexity of scheduling multiple participants at the same time to adhere to randomization across the four conditions. As a result, the study was divided into two separate experiments in which each focused on a single independent variable.

Experiment 1 was conducted with 66 participants randomly assigned as individuals to one of two conditions: (a) with self-explanation prompts and (b) without self-explanation prompts. Experiment 2 included 20 participants who were placed in 10 dyads and then randomly assigned to one of two conditions: (a) with self-explanation prompts and (b) without self-explanation prompts. Across the two experiments, the learning environment, measurements, and procedures were identical.

CHAPTER 2

Method

Experiment 1

Participants and design. Participants were recruited from university physics classes across multiple semesters and received $20 as a token of appreciation for their participation. Sixty-six participants were randomly assigned in equal numbers (*n* = 33) to one of two conditions: (a) with self-explanation prompts and (b) without self-explanation prompts. Each participant was given a unique identifier to match the pretest, posttest, and work booklet that contained each problem in the tutoring session. All the participants were within the age range of 18 to 27 years old with the median age of participants was 20 years old. The sample consisted of 46 males and 20 females. Table 1 summarizes the percentage of participants according to their majors.

Table 1

Summary of Participants according to their Self-reported Majors in Experiment 1

Major	No. of participants	Percentage of Experiment 1 participants
Physics	9	13.4%
Mechanical engineering	6	9.0%
Chemical engineering	5	7.5%
Master's in construction engineering	5	7.5%
Electrical engineering	4	6.0%
Biomedical engineering	4	6.0%
Bachelor's (general)	4	6.0%
Aerospace engineering	3	4.5%
Engineering management	3	4.5%
Master's (general)	3	4.5%
Computer science	3	4.5%
Civil engineering	2	3.0%
Chemistry	2	3.0%
Computer systems engineering	2	3.0%
Biochemistry	2	3.0%
BS geological science	1	1.5%
Engineering	1	1.5%
Construction engineering	1	1.5%
Master's in chemical engineering	1	1.5%
Biophysics	1	1.5%
Master of urban and environmental planning	1	1.5%
Interdisciplinary studies	1	1.5%
Industrial engineering	1	1.5%
Master's in counseling	1	1.5%

Learning environment and materials. The learning environment for this experiment consisted of a video capturing a tutoring dialogue between an experienced tutor and a novice student discussing Newtonian Laws of Motion, which was approximately 45 minutes in length. A screen shot of the learning environment is shown in Figure 1.

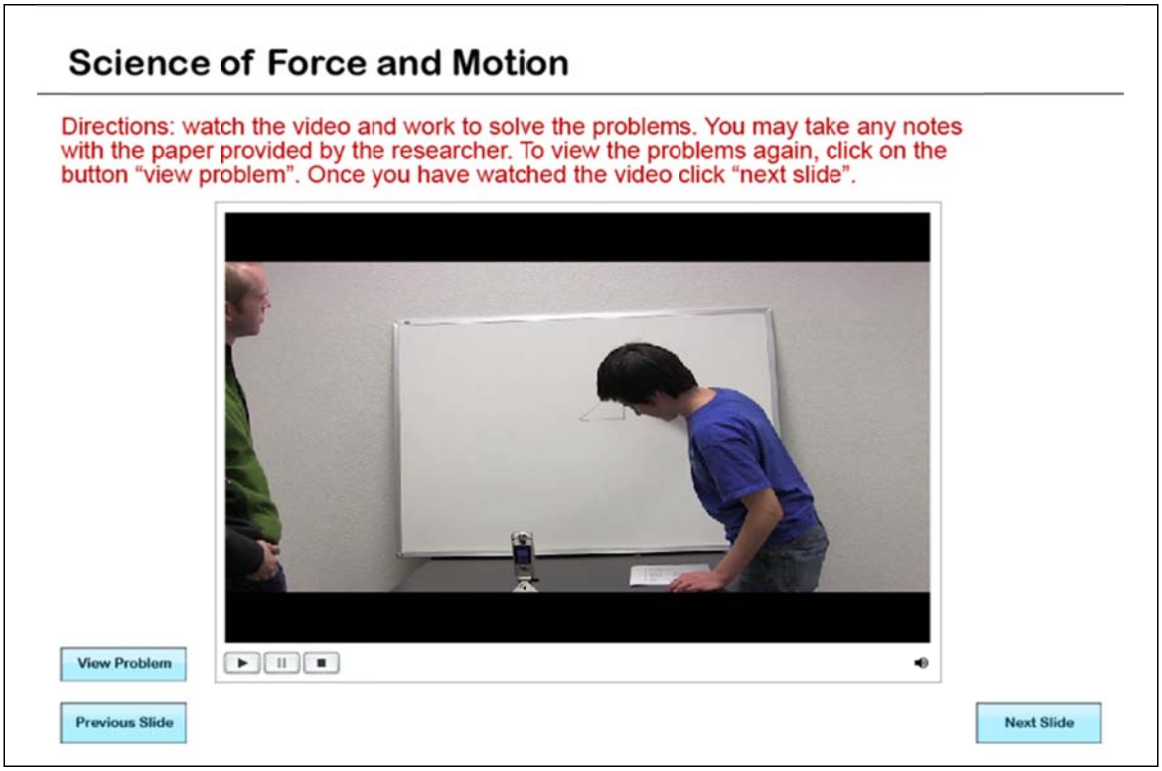


Figure 1. Screenshot of tutoring segment in learning environment. This screenshot illustrates how every tutoring segment in the learning environment was used for both studies.

The professional tutor featured in the video was graduate student with previous experience as a physics instructor, whereas the student was enrolled in an introductory physics class. The problems in the tutoring session related to the Newtonian Laws of Motion, specifically acceleration, force, and tension. The tutoring items are presented in Appendix A. The original professional tutor/tutee session was videotaped and then edited into seven different segments, to give the participants time to answer each prompt. Between each segment, the tutoring problem was presented to the learner. All the participants viewed the identical video segments. The participants were provided controls to pause, stop, or rewind each segment.

The materials and equipment required for this study included *Adobe Captivate*, pretest, posttest, usability survey, and IMI survey. The learning environment was constructed with Adobe Captivate. There were two versions of the learning environment: One version did not include self-explanation prompts, whereas the other version included self-explanation prompts (see Figure 2). The types of self-explanation prompts used in the learning environment were justification prompts and step-focused prompts. Both versions of the learning environment contained identical items and video segments.

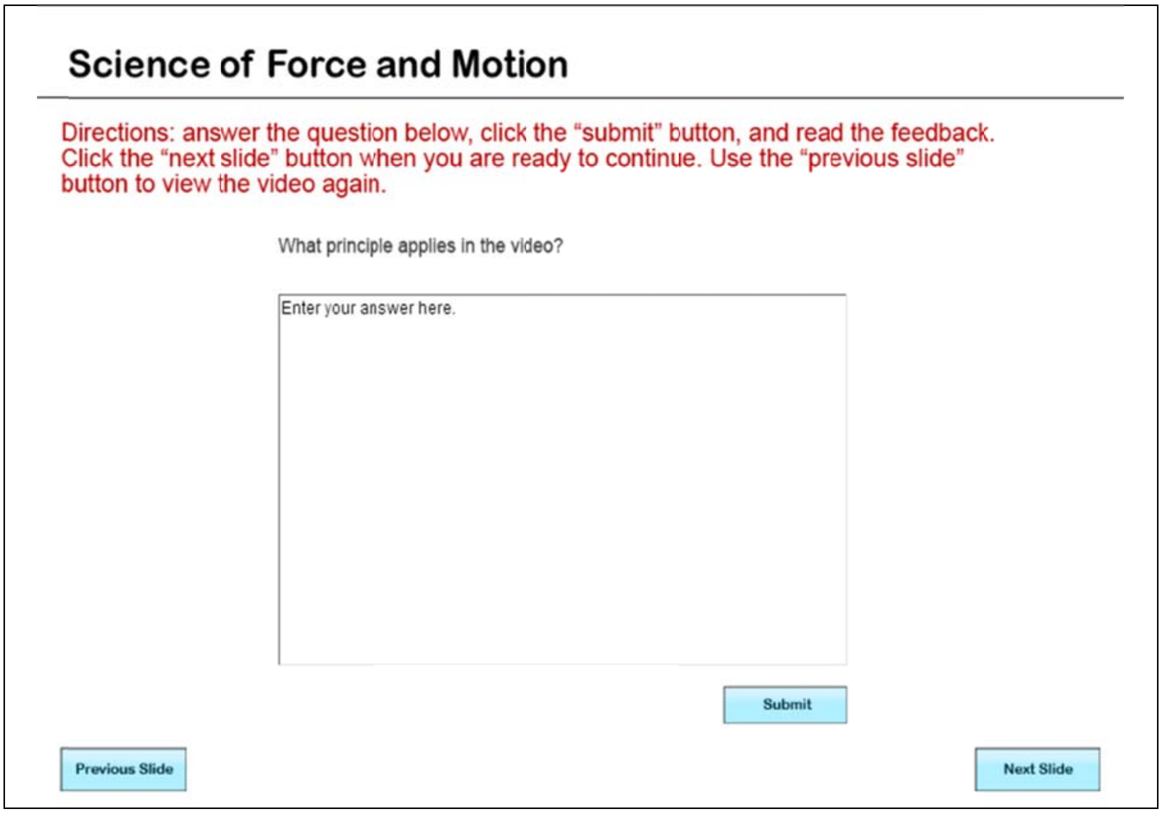


Figure 2. Self-explanation prompt in the learning environment. This screenshot illustrates how self-explanation prompts were presented to the learner in the learning environment used for both studies.

The self-explanation prompts were present on every slide that followed the segment of the video (see Figure 2 for a screenshot of the presentation of the self-explanation prompt). The self-explanations were a mixture of justification and step-focused prompts. Justification and step-focused prompts are both gap filling-prompts, in which learners was expected fill in the missing information. Justification prompts required learners to substantiate problem-solving steps; thus, a learner must “focus her or his processing on the underlying concepts and application conditions for the step” (Nokes et al., 2011, p. 647). For instance, the justification prompt “What principle is being applied here” was used in this study. Step-focused prompts are intended to elicit explanations of specific steps. Nokes et al. (2011) noted that step-focused prompts can provide opportunities for learners to create inference. Within this study, the step-focused prompt, “What is the direction of the Normal Force acting on the block? What is the value of the Normal force when (a) the string is taut, (b) the string is cute loose and the box is sliding down” was used. The self-explanation prompts are presented in Table 2. The response was then recorded in an e-mail and sent to the researcher. The justification and step-focused prompts were chosen in consultation with a subject matter expert in physics. All of the prompt responses were recorded.

Table 2

Justification and Set-focused Self-explanation Prompts Used in the Learning Environment

Self-explanation prompt	Justification or Step focused
• What principle applies here?	Justification
• What is the direction of the Normal Force acting on the block? What is the value of the Normal force when (a) the string is taut, (b) the string is cute loose and the box is sliding down?	Step-focused
• How does acceleration relate to net force?	Justification
• Why is there a "normal force" on the crate? What would the acceleration of the crate be if we excluded it?	Step-focused
• How many principles (Laws of Physics/Mechanics) are applied to solve the pulley problem?	Justification
• How is the acceleration of the blocks related to the mass of each?	Justification

Measures. A pretest was used to measure the prior knowledge of Newtonian mechanics. The pretest was composed of three closed-ended items, two of which consisted of multiple parts (see Appendix B). The items were adapted from Chi et al. (2008), which were, in turn, selected out of a classic physics textbook *Fundamentals of Physics* (Halliday & Resnick, 1981). The items used covered the topics of the Newtonian Laws of Force and Motion. All the problems in the pretest were near-transfer items. The participants were instructed to show their work on a separate sheet of paper.

In contrast, the posttest items were created to measure learning after viewing the tutored problems. Problems in the posttest included concepts discussed in the tutoring videos, which resulted in multiple-part items. Additionally, the multiple-part items helped to determine the quality of the self-explanation answers. In total, the posttest contained four multiple-part items (see Appendix C). As was the case in the pretest, items 1-3 of the posttest required learners to utilize identical Newtonian Mechanics concepts and formulas. The value of the Newtonian Mechanics variables differed from the pretest to posttest, and Item 4 of the posttest did not correspond to any problem in the pretest (Chi et al., 2008); instead, it represented the final question in the tutoring session. Ultimately, Item 4 was omitted from the final analysis of Chi et al. (2008). Therefore, the pretest in the current study consisted of three items, whereas the posttest consisted of three multiple-part items. The pretest and posttest were used to assess learning gains.

All participants completed a demographic questionnaire, in which gender, grade level, current degree, and current age were recorded. Additionally, the demographic survey asked participants to list any Physics courses enrolled in at the university. The demographic questionnaire is presented in Appendix D.

The *System Usability Scale* (Brooke, 1986) was used to assess the overall usability of the learning environment (see Appendix E). The SUS was made up of 10 subjective statements that participants responded to on a Likert scale (1 = *strongly disagree* to 5 = *strongly agree*). The statements included “I think that I would like to use this learning system frequently”, “I found the learning module unnecessarily complex”, and “I though the learning module was easy to use”. The scale requires participants to rate their level of agreement on 10 statements, of which half are positive statements and half are negative statements.

The *Intrinsic Motivation Inventory* was used to measure the motivation of each participant (see Appendix F). The IMI was included in this study to assess the subjective experiences of the participant. The IMI consisted of 45 items that measured these: (a) interest and enjoyment, (b) perceived competence, (c) effort, (d) value, (e) pressure and tension, and (f) choice while performing an activity. The IMI is made up of subjective statements that participants responded to on a Likert scale (1 = *strongly disagree* to 7 = *strongly agree*; Ryan & Deci, 2000).

Procedure. The experiments were carried out in a university computer lab that included nine individual computers and space for pair of participants to work. The participants first completed the pretest. Once it was completed, the participants were instructed to view each segment of the tutoring video and follow the on-screen instructions. Participants were also told they could raise their hand anytime to ask a question. Participants who asked questions were referred to the instructions on the screen (see Figure 2). At the start of every condition, participants were given a work booklet with the pretest, tutoring items, and posttest; the participants were instructed to show their work throughout the module.

After the instruction concluded, each participant completed the questionnaire, the SUS, and the IMI. After completing all the surveys, each participate was informally interviewed. Among the questions that were asked during the interview included “Do you think you would use this type of learning environment for any other subject matter?” and “At any point, did you have an aha-moment?”

Individual students in the observer/self-explanation prompts condition were instructed to fill out an answer to a prompt after each videotaped tutoring segment. The response was then recorded in an e-mail and sent to the researcher.

Students in the individual observer/no self-explanation prompts condition did not fill out any prompts, but were told to follow along with the videos and work out the problems. Appendix G contains a summary of the 25 steps followed in the Experiment 1 and 2; the only difference is Experiment 1 included a single observer, whereas Experiment 2 included dyad observers.

Scoring. Tests were scored to assess learning gains. Each item on the pretest and posttest was scored 0 = *incorrect* or 1 = *correct*. A total percentage correct was calculated for both tests by totaling the score correct and dividing by the number of items on the instrument. The total number of questions on the pretest, which included multiple-part questions, was 5. The total number of questions on the posttest, which included multiple-part question, was 11.

The Systems Usability Scale was scored according to the directions provided by Brooke (1996). The SUS score was calculated by first subtracting 1 from the participant's score for items: 1, 3, 5, 7, and 9, and then subtracting five from the participant's score for items 2, 4, 6, 8, and 10. Afterwards, this number was multiplied by 2.5.

The IMI consisted of seven subscales: interest/enjoyment, perceived competence, effort/importance, pressure/tension, perceived choice, value/usefulness, and relatedness. The total number of items was 45, of which 16 items were reversed scored; to calculate the reverse score, the item score was subtracted from 8. The total possible score for each subscale is presented in Table 3.

Table 3

Possible Scores for Intrinsic Motivation Inventory Subscales

Subscale		Subscale possible score
Interest/enjoyment		33
Perceived competence		34
Effort/importance		19
Pressure/tension		19
Perceived choice		9
Value/usefulness		49
Relatedness		24
	Total	187

Experiment 2

Participants and design. Twenty participants were placed into 10 dyads and then randomly assigned to conditions: (a) pair collaboratively observing with self-explanation prompts and (b) pair collaboratively observing without self-explanation prompts. Each participant was given a unique identifier to match the pretest, posttest, and work. All the participants were within the age range of 18 to 25 years old; the median age of participants was 19 years old. The number of male participants was 18 and the number of female participants was 2. Table 4 summarizes the percentage of participants according to their major.

Table 4

Summary of Participants according to their Self-reported Majors in Experiment 2

Major	No. of participants	Percentage of Experiment 2 participants
Mechanical engineering	5	25%
Chemical engineering	3	15%
Aerospace engineering	2	10%
Biomedical engineering	2	10%
Post-baccalaureate (general)	1	5%
Master's (general)	1	5%
Materials science and engineering	1	5%
N/A	1	5%
Computer science	1	5%
Informatics	1	5%
Bachelor of sciences in physics and computational mathematical sciences	1	5%
Physics	1	5%

Learning environment. The learning environment in Experiment 2 was identical to the learning environment in Experiment 1; however, participants were paired into dyads. Participants in the pair of observers with self-explanation prompts condition were instructed to work together to answer the prompt.

Measures. The materials and instruments used in Experiment 2 were identical to those used in Experiment 1.

Procedure. The procedures in Experiment 2 were identical to the procedures in Experiment 1; however, participants were paired into dyads in Experiment 2. Therefore, the conditions in Experiment 2 were as follows: (a) pair of observers with self-explanation prompts and (b) pair of observers without self-explanation prompts. A portion the dyad groups were video-taped and informally interviewed.

Scoring. The procedures to score the data from the measurements and instruments were identical to those in Experiment 1.

CHAPTER 3

Results

Experiment 1

The first experiment included two conditions: (a) individual observer with self-explanation prompts and (b) individual observer without self-explanation prompts. An independent t-test was conducted to determine whether there was a significant difference in mean scores on learning gains. The means and standard deviations for Experiment 1 are shown in Table 5.

Table 5

Learning Pretest and Posttest Means and Standard Deviations (and Posttest Means and Standard Deviations by Condition) for Experiment 1

	M (*SD*)
Overall Pretest in Experiment 1	33.85 (29.40)
Overall Posttest in Experiment 1	61.12 (19.17)
Condition 1: Individual with self-explanation prompts posttest	61.12 (19.16)
Condition 2: Individual with no self-explanation prompts posttest	64.88 (18.63)

Note. Total possible score on Pretest and Posttest was 100.

Research Question 1. Research Question 1 asked whether self-explanation prompts helped support students' learning while watching a recorded tutorial dialogue. An independent-samples *t*-test was conducted to test the mean difference between individual with self-explanation prompts condition and individual with no self-explanation prompts condition. Results from an independent *t*-test found no significant effect, $t(64) = -.81$, $p = .42$, on the posttest. See Table 6 for the results.

Table 6

Independent t*-test Results (Equal Variances Assumed) for Experiment 1*

	t	*df*	*p* (two-tailed)	*M* (*SD*)
Posttest score	-.808	64	.422	
Condition 1: Individual with self-explanation prompts				61.12 (19.16)
Condition 2: Individual with no self-explanation prompts				64.87 (18.63)

System Usability Scale primary analysis results. An independent-samples *t*-test was conducted to test the difference between the average of the participant scores on the SUS. Results from the independent t-test found no significant effect, $t(64) = 1.58$, $p = .12$. The independent samples *t*-test results for the average SUS scores, as well as means and standard deviations by condition, are presented in Table 7.

Table 7

The Means and Standard Deviations by Condition for System Usability Scale Scores in Experiment 1

	M (*SD*)
Condition 1: Individual with self-explanation prompts	43.12 (19.16)
Condition 2: Individual with no self-explanation prompts	46.88 (18.63)

Note. SUS = System Usability Scale. Total possible score was 100.

Intrinsic Motivation Inventory primary analysis results. An independent-samples *t*-test was conducted was conducted to evaluate the mean difference between the two conditions on the IMI scale. There was no significant effect, $t(64) = 1.86$, $p = .07$. The independent *t*-test results for the average IMI scores, as well as means and standard deviations by condition, are presented in Table 8.

Table 8

The Means and Standard Deviations by Condition for Intrinsic Motivation Inventory Scores in Experiment 1

	M (*SD*)
Condition 1: Individual with self-explanation prompts	73.07 (14.50)
Condition 2: Individual with no self-explanation prompts	65.95 (16.48)

Note. IMI = Intrinsic Motivation Inventory. Total possible score was 187.

Experiment 1 supplemental analyses. An additional test was run to explore learning gain in general regardless of condition between the pretest and posttest. The supplemental analysis for Experiment 1 was performed to look more closely at potential differences in usability of the learning environment across conditions by testing each item on the SUS scale independently. Additionally, an analysis of the mean difference between the two conditions was carried out on each of the seven factors of the IMI.

Paired-samples t-test to evaluate learning gains in Experiment 1. A paired-samples *t*-test was conducted to evaluate learning in general from pretest and posttest regardless of condition. The results indicated that the mean score for the posttest was significantly greater than the mean score for the pretest (M = -29.03, SD = 31.14), $p <$ .01. The 95% confidence interval fell between -36.69 and -21.38. The independent *t*-test results are presented in Table 9.

Table 9

Paired-samples t*-test Results (Equal Variances Assumed) for Experiment 1*

	M (*SD*)	*SE*	95% confidence interval		*t*	*df*	*p* (two-tailed)
			LCL	UCL			
Pretest - Posttest	-29.03 (31.14)	3.83	-36.69	-21.38	-7.57	65	< .01

Note. LCI = Lower confidence limit; UCI = Upper confidence limit.

System Usability Scale supplemental analysis results. An independent-samples *t*-test was conducted to evaluate the mean difference between conditions of the items on the SUS scale; the SUS scale included a total of 10 items. Differences on the scale were significant. The item was “I need to learn a lot of things before I could get going with this learning module.” The significant differences in scores on this item could indicate a flaw in the learning environment; however, these results should be interpreted cautiously due to an increased Type 1 Error. The independent samples *t*-test results by item are presented in Table 10.

Table 10

Independent t*-test Results (Equal Variances Assumed) for System Usability Scale Items for Experiment 1*

	t	*df*	*p* (two-tailed)
1. I think that I would like to use this learning system frequently	1.22	64	.23
2. I think I would need the support of a technical person to be able to use this system	1.90	64	.63
3. I thought the learning module was easy to use	.67	64	.51
4. I found the learning module unnecessarily complex	1.90	64	.58
5. I found the various functions in this learning module were well integrated	-.34	64	.73
6. I thought this system was too inconsistent	.-.17	64	.86
7. I would imagine that most people would learn to use this system very quickly	-.02	64	.98
8. I found the system very cumbersome to use	-1.29	64	.20
9. I felt very confident using the learning module	.38	64	.70
10. I need to learn a lot of things before I could get going with this learning module	-2.37	64	.02

Note. SUS = System Usability Scale.

Intrinsic Motivation Inventory supplemental analysis results. An independent-samples *t*-test was conducted to evaluate the mean differences between the two conditions on the subscale of these: (a) interest, (b) perceived competence, (c) effort/importance, (d) pressure/tension, (e) perceived choice, (f) value/usefulness, and (g) relatedness. The mean on the subscale for Perceived Competence was significantly different in the conditions, $t(64) = 2.83$, $p = .01$. The independent samples *t*-test results by factors are presented in Table 11.

Table 11

Independent-samples t-*test Results (Equal Variances Assumed) for Intrinsic Motivation Inventory Factors for Experiment 1*

IMI Factor	*t*	*df*	*p* (two-tailed)
Interest/Enjoyment	-.13	64	.90
Perceived competence	2.83	64	.01
Effort/Importance	.40	64	.69
Pressure/Tension	.56	64	.58
Value/Usefulness	1.14	64	.26
Relatedness	.73	64	.47
Perceived choice	1.27	64	.21

Note. IMI = Intrinsic Motivation Inventory.

Interpretation of supplemental analysis for Experiment 1. A majority of participants in Experiment 1 appeared to learn from the pretest to posttest according to the paired-samples *t*-test in Table 9. The significance of the item "I need to learn a lot of things before I could get going with this learning module" could indicate that learners struggled with the learning environment; however, according to the independent-samples *t*-test on the IMI scale, most learners felt competent with the material.

Experiment 2

The second experiment included two conditions: (a) paired participants with no self-explanation prompts and (b) paired participants with self-explanation prompts. An independent t-test was conducted to determine whether there was a significant difference in mean scores on learning gains, usability, and motivation. The measurements included a pretest, posttest, Systems Usability Survey, and Intrinsic Motivation Survey (Brooke, 1996; Ryan & Deci, 2000). The means and standard deviations for Experiment 2, as well as the means and standard deviations by condition, are shown in Table 12.

Table 12

Learning Pretest and Posttest Means and Standard Deviations (and Posttest Means and Standard Deviations by Condition) for Experiment 2

	M (*SD*)
Overall Pretest	41.25 (27.24)
Overall Posttest	56.05 (20.86)
Condition 1: Participant pair with self-explanation prompts	49.60 (19.48)
Condition 2: Participant pair with no self-explanation prompts	62.50 (21.13)

Note. Total possible score on Pretest and Posttest was 100.

Research Question 2. Research Question 2 sought to examine what effect does collaboratively observing a tutorial dialogue with self-explanation prompts have on learning. An independent samples t-test was conducted to test the mean difference on the posttest. Results from an independent t-test found no significant effect, $t(18) = 1.42$, $p = .17$, on the posttest.

System Usability Scale primary analysis results. An independent-samples *t*-test was conducted to test the difference between the average of participant scores on the SUS. Results from the independent t-test found no significant effect, $t(18) =-.84$, $p =.41$. The independent samples *t*-test results, as well as means and standard deviations by condition, are presented in Table 13.

Table 13

Independent t*-test Results (Equal Variances Assumed) for Average System Usability Scale Scores for Experiment 2*

	M (*SD*)
Condition 1: Participant pair with self-explanation prompts	26.30 (5.20)
Condition2: Participant pair with no self-explanation prompts	28.68 (1.30)

Note. SUS = System Usability Scale. Total possible score was 100.

Intrinsic Motivation Inventory primary analysis results. An independent-samples t-test was conducted was conducted to evaluate the mean difference between the two conditions on the IMI scale. There was no significant effect, $t(18) = -.87$, $p = .40$. The independent *t*-test is presented in Table 14.

Table 14

Independent t*-test Results (Equal Variances Assumed) for Average Intrinsic Motivation Inventory Scores for Experiment 2*

	M (*SD*)
Condition 1: Participant pair with self-explanation prompts	65.83 (23.61)
Condition 2: Participant pair with no self-explanation prompts	73.52 (15.25)

Note. IMI = Intrinsic Motivation Inventory. Total possible score was 187.

Experiment 2 supplemental analyses. An additional test was run to explore learning gain in general regardless of condition between the pretest and posttest. Similar to the supplemental analysis for Experiment 1, supplemental analyses were performed to look more closely at potential differences in usability of the learning environment across conditions by testing each item on the SUS scale independently. Additionally, an analysis of the mean difference between the two conditions was carried out on each of the seven factors of the IMI.

Paired-samples t-test to evaluate learning gains in Experiment 2. A paired-samples *t*-test was conducted to evaluate learning in general from the pretest and posttest regardless of condition. The results indicated that the mean score for the posttest was significantly greater than the mean score for the pretest. The 95% confidence interval did not contain the value zero and ranged -46.78 to -18.82. The paired-samples *t*-test result is shown in Table 15.

Table 15

Paired-samples t*-test Results (Equal Variances Assumed) for Experiment 2*

	M (*SD*)	*SE*	95% confidence interval		*t*	*df*	*p* (two-tailed)
			LCL	UCL			
Pretest – Posttest	-32.80 (29.86)	6.68	-46.78	-18.82	-4.91	19	< .01

Note. LCI = Lower confidence limit; UCI = Upper confidence limit.

System Usability Scale supplemental analysis results. An independent-samples *t*-test was conducted to evaluate the mean differences on the SUS scale; the SUS scale included a total of 10 items. Differences on multiple items on the scale were not significant; however, these results should be interpreted cautiously due to an increased Type 1 Error. The independent samples *t*-test is presented in Table 16.

Table 16

Independent t*-test Results (Equal Variances Assumed) for System Usability Scale Items for Experiment 2*

	t	*df*	*p* (two-tailed)
1. I think that I would like to use this learning system frequently	1.73	18	.10
2. I think I would need the support of a technical person to be able to use this system	.46	18	.65
3. I thought the learning module was easy to use	.82	18	.42
4. I found the learning module unnecessarily complex	1.25	18	.23
5. I found the various functions in this learning module were well integrated	-.23	18	.82
6. I thought this system was too inconsistent	-1.36	18	.19
7. I would imagine that most people would learn to use this system very quickly	-.91	18	.37
8. I found the system very cumbersome to use	1.17	18	.26
9. I felt very confident using the learning module	-.18	18	.86
10. I need to learn a lot of things before I could get going with this learning module	1.31	18	.21

Note. SUS = System Usability Scale.

Intrinsic Motivation Inventory supplemental analysis results. An independent-samples *t*-test was conducted was conducted to evaluate the mean difference between the two conditions on (a) interest, (b) perceived competence, (c) effort/importance, (d) pressure/tension, (e) perceived choice, (f) value/usefulness, and (g) relatedness. The non-significant test results are presented in Table 17.

Table 17

Non-significant Independent t*-test Supplemental Analysis Results (Equal Variances Assumed) for Intrinsic Motivation Inventory Factors for Experiment 2*

IMI Factor	*t*	*df*	*p* (two-tailed)
Interest/Enjoyment	-.77	18	.45
Perceived competence	-1.60	18	.13
Effort/Importance	-.68	18	.51
Pressure/Tension	-1.79	18	.09
Value/Usefulness	-.06	18	.95
Relatedness	.30	18	.77
Perceived choice	.87	18	.40

Interpretation of supplemental analysis for Experiment 2. A majority of participants in Experiment 2 did learn from the pretest to posttest according to the paired-samples *t*-test in Table 16. The lack of significant differences in the SUS and the Intrinsic Inventory Scale could needs to be considered cautiously, due to the small sample size.

CHAPTER 4

Conclusions and Discussion

Experiment 1

The results for Experiment 1 were not significant across any of the measures. According to the data, self-explanation prompts did not influence learning gains when observing a tutorial dialogue. This finding was contrary to the positive learning gains associated with self-explanation prompts in the research literature, as well as the learning gains associated with the research on the ICAP Framework.

The process of self-explaining is constructive; knowledge is created through providing an explanation for each step of action in problem-solving. Through using prompts, educational technologists and instructors can increase metacognition in learners (Chi et al., 1989). Lin and Atkinson (2013) found that instructional designers can create effective learning environments in which animated visuals are combined with self-explanation prompts.

Nokes et al. (2011) asserts that both justification and step-focused prompts are gap-filling prompts that can be used to solicit missing information in examples in instruction. Moreover, justification prompts can be used to focus on underlying concepts, whereas step-focused prompts encourage the learner to explain each step.

Notably, there are several changes that could have been made to the design of the self-explanation prompts within the learning environment in both Experiment 1 and Experiment 2: (a) The prompt response could have been scaffolded or recorded in such a way as to reduce the impact on working memory, (b) learners could have been given more time to respond to the self-explanation prompt, and (c) the self-explanation prompt could have been presented earlier in the tutoring video. There several ways these changes could have been made: first, in other studies, a *talk-out-loud* procedure was used to record the participants' self-explanations. This procedure would have required more training for the participants, but it would reduce the pressure of typing out a prompt answer. Second, the amount of time between the prompt and giving a response could be increased by editing the learning environment. Last, the location of the prompt within the tutoring video could be manipulated.

Although no overall significant differences were found in the SUS in Experiment 1. A significant difference was found on the item: (a) I need to learn a lot of things before I could get going with this learning module. These results must be interpreted with caution given the inflation of Type 1 Error rate that resulted from the relatively high number of t-tests conducted. The condition that included self-explanation prompts required participants to type in their answer to the prompt and click submit, before moving onto the next tutorial video. This extra step can be seen in Figure 2. The directions are visible and easy to read; however, the processes of self-explaining, typing the answer, and clicking submit could have dramatically increased cognitive load (Sweller, 1988). According to Paas, Renkl, and Sweller (2003) the process of typing in an answer could be viewed as extraneous load. Additionally, some amount of working memory could have been taken up by having to think of the response and then type it out (Sweller, 1988). Consequently, the extraneous load imposed by the learning environment could have led to the significant difference in the three items on the SUS.

Previously, the literature has described the activity of collaboratively observing a tutorial dialogue as constructive or interactive activity within the ICAP framework.

Moreover, activities can be categorized as (a) engaging activities, (b) self-constructive activities, or (c) guided construction activities in instructional dialogue (Chi, 2009).

Experiment 2

The results of Experiment 2 were not significant. According to the data, self-explanation prompts did not influence learners collaboratively observing a tutorial dialogue. This is incongruous with the current literature pertaining to the ICAP Framework. According to the ICAP Framework, interactions with a peer, while observing a dialogue, can be classified as constructive (Chi, 2009). Moreover, the use of self-explanation prompts creates a guided activity in which participants respond; however, there are several possibilities that led to the learner performance (Renkl, 1997).

Within Experiment 2, on the SUS differences in scores were not found. As with Experiment 1, these results must be interpreted with caution given the inflation of Type 1 Error rate that resulted from the relatively high number of *t*-tests conducted. Furthermore, given the small sample, size the findings on the SUS must be interpreted with caution. As in Experiment 1, the learning environment could have exacerbated extrinsic cognitive load, while lowering germane and intrinsic cognitive load; additionally, having to converse and interact with a partner when observing a dialogue could create additional working memory constraints.

Limitations

The current study was originally designed as a 2x2 between-subjects design, with the independent variables being the presence of self-explanation prompts and collaborative learning. The purpose of the 2x2 between-subjects design was to determine whether an interaction between the presence of self-explanation prompts and collaboration produced an interaction; however, given an error in randomization, the proposed 2x2 between-subjects design was split into two experiments. The error in randomization occurred due to conflicts in scheduling participants. This limited the scope of the research questions and prevented the researcher from determining whether an interaction was present between the presence of self-explanation prompts and the number of participants viewing the tutorial dialogue.

After reviewing the learning environment, it is evident that there are concerns that must be addressed. The first concern is the quality of the video recordings of the tutoring session. In some moments, throughout the video, the tutor and tutee construct notes on a whiteboard; several participants commented that the writing on the whiteboard was difficult to read. Although the tutor and tutee went over each item step-by-step, the notes on the whiteboard might have been difficult to read in some areas.

In each experiment, the self-explanation conditions required participants to type in their answer and then hit a submit button. Although instructions on submitting the answer was provided for every prompt answer, this necessitated the need for an entire slide devoted to answering the prompt; each tutoring segment was followed by an additional slide describing the prompt and instructions. The additional slides might have caused an interruption in the flow of the tutorial dialogue, thus limiting the participants thought process.

Future Research

At the beginning of this study, one of the main purposes was to investigate the presence of an interaction between self-explanation prompts and the number of observers; however, the study had to be split into two experiments. To investigate the presence of an interaction a two-tiered randomization method could be employed. This method of randomization would have participants first randomly assigned to individual observer or pair of observes, and then the participants could be randomly assigned to either with self-explanation prompts or without self-explanation prompts.

When creating the learning environment for future research each part of the example and solution to the tutoring problem needs to be created in a way in which the work of the tutor and tutee are clean and visible. This can be accomplished by creating each worked-out step of the problem on a poster board. The poster board can be displayed while the tutor and tutee work through the problem. This would allow the researchers to verify the readability of the work of the tutor and tutee. Another way to increase readability would be to create a computer image of the work that takes place on the whiteboard.

To clearly capture participants' responses to each self-explanation prompt, a talk-out-loud protocol needs to be developed. This talk-out-loud protocol would be similar to those used in Chi et al.'s (2008) studies and others. Through using a talk-out-loud protocol, the self-explanation prompts could be answered without interrupting the flow of the tutorial dialogue and thus impacting cognitive load. Additionally, a recording of the talk-out-loud interactions between pairs would more clearly represent the response of a self-explanation prompt, because participants would not be required to type their answer.

Due to the requirement of participants to record their answers to prompts in a dialogue box, the entire 45-minute tutoring dialogue had to be segmented into 7 different sections. This segmentation could have disrupted the natural flow of the tutorial dialogue or led to extraneous cognitive load. In future iterations, the tutorial dialogue will not be segmented. This can be accomplished by the use of a talk-out-loud protocol and recording either the audio or video.

All the prompt responses were recorded in the current study, additionally dyad groups were recorded in Experiment 2; it could prove valuable to analyze the prompt responses and recorded dialog groups. Through this analysis the self-explanation prompts could be refined and any interactions between participants could be analyzed.

In sum, Experiments 1 and 2 did not find evidence to support the use of self-explanation prompts when observing a tutorial dialogue; however the supplemental analysis in Experiment 1 revealed significant findings. The independent-samples t-test conducted on the SUS in Experiment 1 showed a mean difference on the item, “I need to learn a lot of things before I could get going with this learning module”; however, this finding should be interpreted cautiously due to a Type 1 Error. Additionally, a significant majority of participants felt that they understood the material and perceived themselves as competent in the material. This finding could be explored in future research through an analysis of the self-explanation prompt answers. The findings in the supplemental analysis of Experiment 2 did not show significance on the SUS or the IMI scale. The low sample size makes it difficult to interpret the results; however, a future analysis of the prompt responses and recorded dyad groups could prove valuable in determining any participant engagement.

References

Adams, D. M., & Clark, D. B. (2014). Integrating self-explanation functionality into a complex game environment: Keeping gaming in motion. *Computers & Education*, *73*, 149-159. https://doi.org/http://dx.doi.org.ezproxy1.lib.asu.edu/10.1016/j.compedu.2014.01.002

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: effects of self-explanation prompts and fading worked-out steps. *Journal of Educational Psychology*, *95*, 774-783. https://doi.org/10.1037/0022-0663.95.4.774

Berthold, K., Eysink, T. H. S., & Renkl, A. (2009). Assisting self-explanation prompts are more effective than open prompts when learning with multiple representations. *Instructional Science*, *37*, 345-363. https://doi.org/10.1007/s11251-008-9051-z

Brooke, J. (1996). SUS: A quick and dirty usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & I. L. McClelland (Eds.), *Usability evaluation in industry* (pp. 189-194). London: Taylor & Francis.

Chi, M. T. H. (2009). Active-Constructive-Interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*, 73-105. https://doi.org/10.1111/j.1756-8765.2008.01005.x

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*, 145-182. https://doi.org/10.1207/s15516709cog1302_1

Chi, M. T. H., Leeuw, N., Chiu, M.-H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*, 439-477. https://doi.org/10.1207/s15516709cog1803_3

Chi, M. T. H., Kang, S., & Yaghmourian, D. L. (2017). Why students learn more from dialogue- than monologue-videos: analyses of peer interactions. *Journal of the Learning Sciences*, *26*, 10-50. https://doi.org/http://dx.doi.org.ezproxy1.lib.asu.edu/10.1080/10508406.2016.1204546

Chi, M. T. H., Roy, M., & Hausmann, R. G. M. (2008). Observing tutorial dialogues collaboratively: insights about human tutoring effectiveness from vicarious learning. *Cognitive Science*, *32*, 301-341. https://doi.org/http://dx.doi.org.ezproxy1.lib.asu.edu/10.1080/03640210701863396

Chi, M. T. H., & VanLehn, K. A. (1991). The content of physics self-explanations. *The Journal of the Learning Sciences*, *1*, 69-105. Retrieved from http://www.jstor.org.ezproxy1.lib.asu.edu/stable/1466657

Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*, 219-243. doi: 10.1080/00461520.2014.965823

Clark, D. B., Virk, S. S., Barnes, J., & Adams, D. M. (2016). Self-explanation and digital games: adaptively increasing abstraction. *Computers & Education*, *103*, 28-43. https://doi.org/http://dx.doi.org.ezproxy1.lib.asu.edu/10.1016/j.compedu.2016.09.010

Crippen, K. J., & Earl, B. L. (2007). The impact of web-based worked examples and self-explanation on performance, problem-solving, and self-efficacy. *Computers & Education*, *49*, 809-821. https://doi.org/http://dx.doi.org.ezproxy1.lib.asu.edu/10.1016/j.compedu.2005.11.018

De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2011). Improved effectiveness of cueing by self-explanations when learning from a complex animation. *Applied Cognitive Psychology*, *25*, 183-194. doi:10.1002/acp.1661

Hilbert, T. S., & Renkl, A. (2009). Learning how to use a computer-based concept-mapping tool: self-explaining examples helps. *Computers in Human Behavior*, *25*, 267-274. https://doi.org/http://dx.doi.org.ezproxy1.lib.asu.edu/10.1016/j.chb.2008.12.006

Hsu, C.-Y., Tsai, C.-C., & Wang, H.-Y. (2016). Exploring the effects of integrating self-explanation into a multi-user game on the acquisition of scientific concepts. *Interactive Learning Environments*, *24*, 844-858. https://doi.org/http://dx.doi.org.ezproxy1.lib.asu.edu/10.1080/10494820.2014.926276

Lin, L., & Atkinson, R. K. (2013). Enhancing learning from different visualizations by self-explanation prompts. *Journal of Educational Computing Research*, *49*, 83-110. https://doi.org/http://dx.doi.org.ezproxy1.lib.asu.edu/10.2190/EC.49.1.d

Muldner, K., Lam, R., & Chi, M. T. (2014). Comparing learning from observing and from human tutoring. *Journal of Educational Psychology, 106*, 69-85. doi: 10.1037/a0034448

Nokes, T. J., Hausmann, R. G. M., VanLehn, K., & Gershman, S. (2011). Testing the instructional fit hypothesis The case of self-explanation prompts. *Instructional Science*, *39*, 645-666. https://doi.org/10.1007/s11251-010-9151-4

O'Neil, H. F., Chung, G. K. W. K., Kerr, D., Vendlinski, T. P., Buschang, R. E., & Mayer, R. E. (2014). Adding self-explanation prompts to an educational computer game. *Computers in Human Behavior*, *30*, 23-28. https://doi.org/http://dx.doi.org.ezproxy1.lib.asu.edu/10.1016/j.chb.2013.07.025

Paas, F., Renkl, A. & Sweller, J. (2003) Cognitive load theory and instructional design: recent developments. *Educational Psychologist, 38*, 1-4. doi:10.1207/S15326985EP3801_1

Rau, M. A., Aleven, V., & Rummel, N. (2015). Successful learning with multiple graphical representations and self-explanation prompts. *Journal of Educational Psychology*, *107*, 30-46. https://doi.org/http://dx.doi.org.ezproxy1.lib.asu.edu/10.1037/a0037211

Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science*, *21*, 1-29. https://doi.org/http://dx.doi.org.ezproxy1.lib.asu.edu/10.1016/S0364-0213(99)80017-2

Renkl, A. (2002). Worked-out examples: Instructional explanations support learning by self-explanations. *Learning and Instruction*, *12*, 529-556. https://doi.org/10.1016/S0959-4752(01)00030-5

Renkl, A., Stark, R., Gruber, H., & Mandl, H. (1998). Learning from worked-out examples: The effects of example variability and elicited self-explanations. *Contemporary Educational Psychology, 23*, 90-108. https://doi.org/10.1006/ceps.1997.0959 https://doi.org/http://dx.doi.org.ezproxy1.lib.asu.edu/10.1006/ceps.1997.0959

Ronchetti, M. (2010). Using video lectures to make teaching more interactive. *International Journal of Emerging Technologies in Learning*, *5*. https://doi.org/10.3991/ijet.v5i2.1156

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*, 68-78. https://doi.org/10.1037/0003-066X.55.1.68

Schworm, S., & Renkl, A. (2006). Computer-supported example-based learning: when instructional explanations reduce self-explanations. *Computers & Education*, *46*, 426-445. doi: 10.1016/j.compedu.2004.08.011

Schworm, S., & Renkl, A. (2007). Learning argumentation skills through the use of prompts for self-explaining examples. *Journal of Educational Psychology*, *99*, 285-296. https://doi.org/http://dx.doi.org.ezproxy1.lib.asu.edu/10.1037/0022-0663.99.2.285

Sweller, J. (1988). Cognitive load during problem-solving: effects on learning. *Cognitive Science, 12,* 257-285. doi: 10.1207/s15516709cog1202_4

Yeh, Y.-F., Chen, M.-C., Hung, P.-H., & Hwang, G.-J. (2010). Optimal self-explanation prompt design in dynamic multi-representational learning environments. *Computers & Education*, *54,* 1089-1100.

APPENDIX A

SYSTEM USABILITY SCALE

System Usability Scale (Brooke, 1996)

System Usability Scale: 1 – 5 likert scale: 1 = strongly disagree, 5= strongly agree

1. I think that I would like to use this learning system frequently.
2. I found the learning module unnecessarily complex.
3. I thought the learning module was easy to use.
4. I think I would need the support of a technical person to be able to use this system.
5. I found the various functions in this learning module were well integrated.
6. I thought this system was too inconsistent
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the learning module.
10. I need to learn a lot of things before I could get going with this learning module.

APPENDIX B

INTRINSIC MOTIVATION INVENTORY

Intrinsic Motivation Inventory (Ryan and Deci, 2000)

1 - 7 Likert rating: 1 = strongly disagree, 7 = strongly agree

Interest/Enjoyment

1. I enjoyed doing this activity very much
2. This activity was fun to do.
3. I thought this was a boring activity.
4. This activity did not hold my attention at all.
5. I would describe this activity as very interesting.
6. I thought this activity was quite enjoyable.
7. While I was doing this activity, I was thinking about how much I enjoyed it.

Perceived Competence

1. I think I am pretty good at this activity.
2. I think I did pretty well at this activity, compared to other students.
3. After working at this activity for a while, I felt pretty competent.
4. I am satisfied with my performance at this task.
5. I was pretty skilled at this activity.
6. This was an activity that I couldn't do very well.

Effort/Importance

1. I put a lot of effort into this.
2. I didn't try very hard to do well at this activity
3. I tried very hard on this activity.
4. It was important to me to do well at this task.
5. I didn't put much energy into this.

Pressure/Tension

1. I did not feel nervous at all while doing this.
2. I felt very tense while doing this activity.
3. I was very relaxed in doing these.
4. I was anxious while working on this task.
5. I felt pressured while doing these.

Perceived Choice

1. I believe I had some choice about doing this activity.
2. I felt like it was not my own choice to do this task.
3. I didn't really have a choice about doing this task.
4. I felt like I had to do this.
5. I did this activity because I had no choice.
6. I did this activity because I wanted to.
7. I did this activity because I had to.

Value/Usefulness

1. I believe this activity could be of some value to me.
2. I think that doing this activity is useful for ______________________
3. I think this is important to do because it can _____________________
4. I would be willing to do this again because it has some value to me.
5. I think doing this activity could help me to _____________________
6. I believe doing this activity could be beneficial to me.
7. I think this is an important activity.

Relatedness

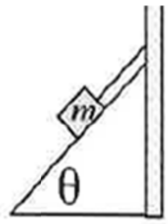
1. I felt really distant to this person.

2. I really doubt that this person and I would ever be friends.
3. I felt like I could really trust this person.
4. I'd like a chance to interact with this person more often.
5. I'd really prefer not to interact with this person in the future.
6. I don't feel like I could really trust this person.
7. It is likely that this person and I could become friends if we interacted a lot.
8. I feel close to this person.

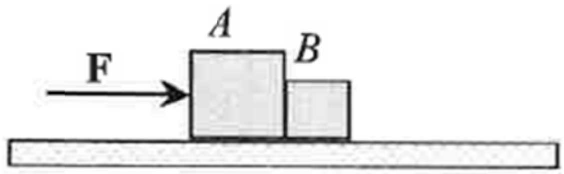
APPENDIX C

ITEMS PRESENTED IN THE TUTORING SESSION

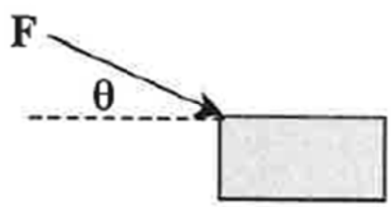
1. A block is attached to a string that is tied to a wall. The block is resting on a smooth plane inclined at an angle 0 with the horizontal.
   a. What is the tension in the string?
   b. What is the force exerted by the wall on the string? Block?
   c. The string is not cut. Find the acceleration of the block if $\Theta=30^0$ and MA=10 kg.
2. Two blocks A and B are in contact with each other on a smooth floor. A force of 10N is applied to the



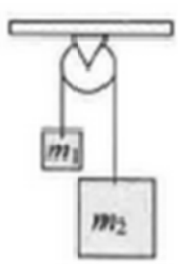
blocks as shown in the figure. Masses of the blocks are 2kg and 3kg respectively.
   a. Find the acceleration of the blocks.
   b. Find net force acting on block B.
   c. Find force exerted by block B on A.



3. A person pushes a crate on a smooth surface. He is applying force at an angle of Θ with the horizontal.
   a. Find the normal reaction of the floor on the crate.
   b. If the mass of the crate is 10 kg, the magnitude of the force is 5N and Θ=300 what will be the acceleration of the crate?
   c. Two blocks, m1 and m2 are connected to two ends of a weightless string which passes over a



weightless frictionless pulley. The pulley is supported from the ceiling.
   d. Find acceleration of the blocks.
   e. Find tension in the string.
   f. What is the force exerted by the pulley support on the ceiling?

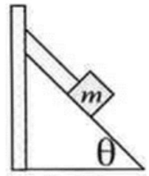


   g. If m_1=4.2 kg, m_2=5.6 kg, how far will m_2 be displaced in 4_s?

APPENDIX D

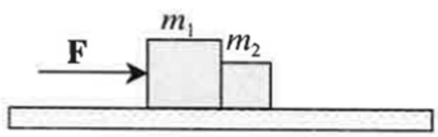
PRETEST ITEMS

1. A block m = 30 kg is attached to a string, which is tied to the wall. The block is resting on a smooth incline at an angle Θ =40 ₒ with the floor. What is the tension in the spring



2. Two blocks are in contact on a frictionless table. A horizontal force is applied to one block.

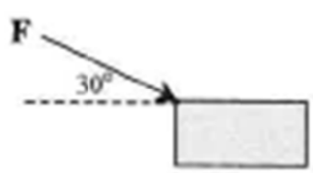
   a. If *ml*= 2.0 kg, *m2* = 1.0 kg, and *F* = 3.0 N, find the force between the two blocks.



(Figure 1)

3. A person pushes a 20kg block at an angle of 8=30° with the floor (see figure below). A force of 20 N is applied to the block. Ignoring friction:

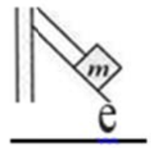
   a. Draw the free body diagram for the block.
   b. Find the normal reaction of the floor on the block.
   c. What is the acceleration of the block?



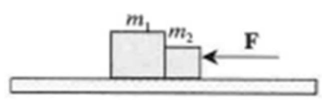
APPENDIX E

POSTTEST ITEMS

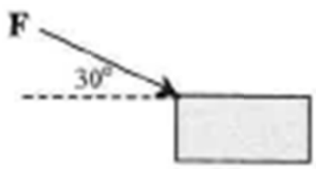
1. A block M =20kg is attached to a string, which is tied to the wall. The block is resting on a smooth incline at Θ=40_0 with the floor.
    a. What is the tension in the string?
    b. What is the force exerted by the wall on the string?
    c. Imagine that the string is cut. Find the acceleration of the block.



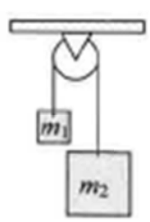
2. Two blocks are in contact on a frictionless table. A horizontal force is applied to one block, as shown below. If m_1=5kg, m_2=4kg, and *F*=20 N:
    a. Find the acceleration of the blocks.
    b. Find the net force acting on the block m_2.
    c. Find the force exerted by Block m_1 on m_2.



3. A person pushes a 20kg block at an angle of Θ=30° with the floor (see figure below). A force of 20 N is applied to the block. Ignoring friction:
    a. Draw the free body diagram for the block.
    b. Find the normal reaction of the floor on the block.
    c. What is the acceleration of the block?



4. Two blocks, m_1=10kg and m_2=20kg are connected by a string that passes over a frictionless and massless pulley.
    a. Find the tension in the string.
    b. Find the acceleration of the masses.



APPENDIX F

DEMOGRAPHIC SURVEY ITEMS

1. What is your gender?
    a. Male
    b. Female
2. What is your current grade level?
    a. Freshman
    b. Sophomore
    c. Junior
    d. Senior
    e. Graduate
3. What is the current degree you are seeking?
4. List any Physics course you are enrolled in at the current university.
5. What is your current age?
6. What condition are you assigned to?

APPENDIX G

SUMMARY OF THE EXPERIMENTS

Step	Action	Prompt
1	Participant sign consent form.	
2	Every participant individually completed the pretest.	
3	Participant read over the problem in video segment 1.	
4	Participant view video segment 1.	
5	Participant answer the self-explanation prompt.	what principle applies in the video?
6	Participant read over the problem in video segment 2.	
7	Participant view video segment 2.	
8	Participant answer self-explanation prompt.	What is the direction of the Normal Force acting on the block? What is the value of the Normal Force when (a) the string is taut (b) the string is cut loose and the box is sliding down?
9	Participant read over the problem in video segment 3.	
10	Participant view video segment 3.	
11	Participant answer the self-explanation prompt	How does acceleration relate to net force?
12	Participant read over the problem in video segment 4	
13	Participant view video segment 4.	
14	Participant answer self-explanation prompt.	Why is there a "normal force" on the crate? What would the acceleration of the crate be if we excluded it?
15	Participant read over the problem in video segment 5.	

16	Participant view video segment 5.	
17	Participant answer self-explanation prompt.	How many principles (laws of physics/mechanics) are applied to solve the pulley problem?
18	Participant read over the problem in video segment 6.	
19	Participant view video segment 6.	
20	Participant answer self-explanation prompt.	Why can we solve for two unknowns if we have two equations that include them? When does this NOT work?
21	Participant reads over the problem in video segment 7.	
22	Participant view video segment 7.	
23	Participant answer the self-explanation prompt.	How is the acceleration of the blocks related to the mass of each?
24	Every participant is given a posttest to be completed individually.	
25	Every participant completes the demographic, usability, and IMI survey.	