Summer 2013

The Effects of Generative Strategies in Instructional Simulations on Learning, Cognitive Load, and Calibration Accuracy

Jennifer R. Morrison
Old Dominion University

THE EFFECTS OF GENERATIVE STRATEGIES IN INSTRUCTIONAL SIMULATIONS ON LEARNING, COGNITIVE LOAD, AND CALIBRATION ACCURACY

by

Jennifer R. Morrison
B.M. May 2001, Central Michigan University
M.S. December 2010, Old Dominion University

A Dissertation Submitted to the Faculty of Old Dominion University in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

EDUCATION

OLD DOMINION UNIVERSITY
August 2013

Approved by:

Ginger S. Watson (Director)

Linda Bol (Member)

Steven M. Ross (Member)

ABSTRACT

THE EFFECTS OF GENERATIVE STRATEGIES IN INSTRUCTIONAL SIMULATIONS ON LEARNING, CALIBRATION ACCURACY, AND COGNITIVE LOAD

Jennifer R. Morrison
Old Dominion University, 2013
Director: Dr. Ginger S. Watson

Instructional simulations can provide a powerful medium for learners to interact with a model representing underlying principles of content or phenomena. While a promising medium for developing a learner's own mental model, reviews of simulation learning have revealed less than promising results (Bangert-Drowns, Kulik, & Kulik, 1985; Kulik & Kulik, 1991), perhaps due to the lack of instructional supports inherent with a discovery-based approach. This study examined the use of generative strategies as an instructional support to promote learning from a physics simulation. Generative strategies, originally proposed by Wittrock (1974, 1989), strengthen understanding by prompting learners to create meaning between new information and prior knowledge or experience. These strategies provide learners with the feedback necessary for reflection in relation to the self-regulatory process described by Zimmerman (2000). Last, engaging in these strategies may direct attention to germane resources necessary for schema construction as described by cognitive load theory (Sweller, Ayres, & Kalyuga, 2011).

Results of this study indicated that principle learning was improved when undergraduate participants paraphrased or predicted and self-explained using a guided discovery approach. Calibration accuracy, by means of predicting anticipated test performance, was also improved for learners engaging in generative strategies as

compared to a control group. Postdiction of test performance indicated a directional trend favoring participants who predicted and self-explained. Test performance was strongly correlated (r=.59) with the thoroughness of generative content between treatment groups and the quality of self-explanations indicated a marked relationship with test performance (r=.78). Generative strategies also led to significant differences in mental effort, assessments of performance, and levels of frustration between treatment groups. Specifically, participants who predicted and self-explained reported significantly higher levels of mental effort than the other two groups. These participants reported decreased levels of confidence than the paraphrase group and higher levels of frustration than the control group. Finally, the incorporation of generative strategies did not influence participants' interest in the instructional content.

Keywords: instructional simulations, principle learning, guided discovery, self-regulation, calibration, cognitive load, mental effort.

(COPYRIGHT PAGE)

This dissertation is dedicated to my parents who let me discover, on my own and in a somewhat round-about manner, a field I am sure they would agree that I was always most suited for.

ACKNOWLEDGEMENTS

I thank my advisor, Dr. Watson, for her attention and mentorship throughout my graduate studies. She immersed me in the literature of our field with my first task as a graduate research assistant, leading to a very strong passion for research and ultimately the pursuit of a doctoral degree. Our numerous conversations fed my constant desire to learn more, and I will always value the time she devoted to me.

I also thank my other committee members, Dr. Bol and Dr. Ross. Dr. Bol provided great (and very quick) feedback on my drafts, introduced me to areas outside of my field, and has become a wonderful mentor. I am grateful for the attention Dr. Ross has given to me over the years, not to mention the contributions and wonderful feedback he has provided on my dissertation.

I would definitely not have made it through this journey if it wasn't for my family. My father entertained numerous questions as I've progressed through my studies, challenging me and also providing guidance. My mother lent wonderful emotional support and tolerated many dinner-table conversations regarding my latest area of research. My brother may have learned much more about my field than he could ever hope (or want), having been a willing test subject, practice presentation observer, and website developer for my stimulus materials. I thank you all so much for your support.

Last, and certainly not least, I thank my friends. They have listened to my rants, late-night discoveries, and have tried to make me take much-needed breaks from my studies. I also thank those classmates that helped to make this study possible, including my content expert, Kate Sliva, simulation developer, Shane Brennan, and content rater, Alan Reid.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Introduction

Instructional simulations provide a powerful medium for learners to interact with the model of a phenomenon and ultimately develop their own mental model to support problem solving and reasoning (Alessi & Trollip, 2001). Interacting with a simulation allows learners to explore the underlying system or phenomenon, adjust variables to observe effects, and to explore realistic and hypothetical situations, without the stress or risk associated with a real-life environment (van Berkum & de Jong, 1991). Simulations are often used to teach principles, where learners explore causal relationships to create a meaningful understanding of the principle represented in the simulation (Reigeluth & Schwartz, 1989).

Learning in a simulation environment is a form of scientific discovery learning (de Jong & van Joolingen, 1998) where learners experience a cycle of planning, executing, and evaluating interactions with the model (Rivers & Vockell, 1987). Scientific discovery learning using a simulation entails learners manipulating variables and observing the effects to induce the characteristics of the underlying model (de Jong & van Joolingen, 1998; Reigeluth & Schwartz, 1989).

Although promising, reviews of computer-enriched instruction including simulation-based instruction indicate small effects on improving learning for secondary students (Bangert-Drowns et al., 1985), moderate improvements at the postsecondary level, and negligible improvements at the precollege level (Kulik & Kulik, 1991) These

conclusions coincide with an assertion by Reigeluth and Schwartz (1989) that the weakest aspect of instructional simulations is the instructional component.

Simulation effectiveness may be limited by problems inherent to the discovery learning process such as creating hypotheses, designing and conducting effective experiments to test hypotheses, making appropriate conclusions from their observations, and issues associated with self-regulating the discovery learning process (de Jong & van Joolingen, 1998). These difficulties are a particular concern for novice learners who do not possess the proper schema to integrate new information with existing knowledge (Tuovinen & Sweller, 1999) and have prompted related research guided by cognitive load theory (Sweller et al., 2011).

While a variety of instructional supports have been developed to assist learners through the discovery process and potentially manage cognitive load, the inclusion of generative strategies may facilitate a deeper processing and integration of content. The generative model of learning and teaching originally proposed by Wittrock (1974, 1989) is founded on knowledge of cognitive processes and research on comprehension, knowledge acquisition, attention, motivation, and transfer (Wittrock, 1992). Generative learning involves making meaningful relations both among concepts and between external information and existing knowledge, resulting in understanding and comprehension. Similar to the levels of processing theory (Craik & Lockhart, 1972), when learners relate new information to existing knowledge, they will process this new information at a deeper level to strengthen memory traces and make the new content more memorable. Generative learning strategies, therefore, encourage learners to activate prior knowledge and stimulate construction of meaningful relations.

The inclusion of generative strategies to prompt learners in making meaningful relations with new information may also assist in self-regulation. Self-regulation refers to learners' active participation in the learning process in regards to behavior, motivation, and metacognition (Zimmerman, 1986). Self-regulated learning is considered a cyclical process consisting of three phases: forethought, performance control, and self-reflection (Zimmerman, 2000). The forethought phase allows for the establishment of learning goals, the development of coinciding plans to obtain goals, and the selection of appropriate learning strategies. The performance phase involves processes that occur during learning, such as engaging in learning strategies and comprehension monitoring. The self-reflection phase occurs after performance where learners evaluate their efforts and adjust future endeavors based on feedback (Zimmerman, 2000). Self-regulated learners are aware of the knowledge they do or do not possess (Zimmerman, 1990), and may be described as being well calibrated, able to make accurate predictions of anticipated performance (Hacker, Bol, & Bahbahani, 2008). Prompting learners to engage in generative learning strategies can facilitate the self-regulatory process by encouraging active construction of knowledge. The products of knowledge construction can provide learners with feedback on progress towards learning goals and assist in calibration accuracy, a monitoring process in self-regulated learning.

Generative strategies may be a powerful aid to support learners' construction of knowledge. However, little research has explored the application of generative strategies within instructional simulations and the effects on working memory resources as described by cognitive load theory (Sweller et al., 2011). Instructional simulations provide an environment for learners to explore content, manipulate variables, and observe

the effects of their actions. Simulations are particularly appropriate for physics, a content area that allows for the exploration of principles. Unfortunately, instructional simulations are often ineffective as they are designed around a discovery-based approach. The purpose of this study is to extend existing research on generative strategies in print and computer-based instruction. Of particular interest is the learning efficacy of generative strategies in a physics simulation environment to promote deeper processing of content, potentially overcome issues associated with cognitive load, and improve learners' self-regulatory processes.

Literature Review

The following literature review is divided into five sections. Research pertaining to principle learning is presented, as this area provides a foundation for a guided discovery approach in simulation instruction. This is followed by a review of research pertaining to the effectiveness of specific generative strategies, paraphrasing and prediction with self-explanation that may be implemented in a simulation environment. The literature review concludes with a discussion on calibration.

Principle Learning

A principle, also referred to as a rule, is a statement of generality that describes a relationship between concepts. Principle instruction may take the form of an expository approach, a statement of the principle with examples, or a discovery-based approach, where several examples are presented and the learner induces the principle (Markle, 1969). A criticism of the discovery-based approach is that it risks learners incorrectly inducing a principle, and therefore it has been suggested to implement this approach once the learner has more experience with the principle (Evans & Homme, 1962). A

compromise between these two approaches is a guided discovery method, where learners receive some sort of guidance or hints to induce a principle from examples provided. In a meta-analysis by Alfieri, Brooks, Aldrich, and Tenenbaum (2011), guided discovery learning was compared to discovery and expository methods across 56 studies. The authors concluded that enhanced discovery, where learners were provided with guidance or feedback throughout the learning process, led to improved learning when compared with other methods across a variety of domains.

Prior research has examined the effects of a guided discovery approach and generally favors this method to a discovery approach for principle learning. For example, differing initial amounts of direction for principle acquisition were provided to undergraduate participants in a study by Craig (1956). The principles, which varied between sets, grouped four out of five words by the sounds of the words, spelling, or a familiar combination. The discovery group was provided with direction that an organizing principle existed, whereas the guided discovery group was provided with a brief general statement that indicated the nature of the relationship. Principle acquisition and retention favored participants in the guided discovery group, who learned significantly more principles than the discovery group, although the two groups performed equally well on a test of transfer to new principles. Results from this study indicate that offering learners guidance in discovering principles is more effective than merely suggesting a principle exists and requiring learners to discover the principle.

Similar results were found by Kittell (1957) who examined principle acquisition with 6^{th} graders. The three groups in this study differed on the amount of instruction regarding a word grouping principle provided: minimum guidance (principle exists),

intermediate (hint of the nature of the principle), and maximum (statement of principle in addition to the correct answer). Performance on the tests of application revealed equivalent performance between the maximum and intermediate groups, both of which performed significantly higher than the minimum, or discovery, group. In contrast to Craig's (1956) study, performance on measures of near and far transfer favored the intermediate group over both the maximum and minimum groups. Additionally, measures of retention over a two- and four-week delay also revealed that participants learning through the intermediate guided discovery approach retained significantly more principle learning than the other two groups. Overall, results of this study coincide with the Craig (1956) study, suggesting that guided discovery, in the form of hints regarding the nature of a principle, is more effective than a discovery approach. Furthermore, presenting the principle along with the correct answer to examples may result in mere rote memorization of the rule and not application of the rule to novel instances.

An additional study by Gagne and Brown (1961) lends support to the use of a guided discovery approach over an expository or discovery approach. The expository group in this study received a statement of the formula after two number series, whereas the discovery group received introductory items followed by a prompt to determine the formula. Participants in the guided discovery group received the same introductory items, followed by a series of prompts that guided the learner in establishing relationships between numbers in the series. After the prompts, the participants were asked to state the rule shown in the number series. All participants then applied the principle to 40 examples, and errors as well as hints used were documented. Although the discovery group required less time during principle acquisition, the guided discovery group

completed the test with the least amount of hints and the shortest amount of time than the discovery and expository groups. The guided discovery method was an effective method for principle learning and acquisition and appears to be a compromise between explicitly providing learners with a principle and allowing learners to discover a principle without support.

Summary and directions. Results of the studies reviewed suggest that a guided discovery approach is more effective when compared to a pure discovery approach with different age groups and on measures of principle acquisition, application, transfer, and retention. In these studies, learners were provided with guidance or hints to correctly induce principles, improving overall learning and retention. What remains unknown from the existing research is how increasing the depth of processing of instruction by means of generative strategies may further facilitate principle acquisition when paired with a guided discovery approach.

Cognitive load theory. One potential explanation for the superiority of a guided discovery approach over a pure discovery approach is described by cognitive load theory (Sweller et al., 2011). Cognitive load theory is based on human cognitive architecture, specifically the limited capacity of working memory and the unlimited capacity of long-term memory (Sweller et al., 2011). This theory has led to the derivation of a number of principles for the design of instructional materials based on three dimensions of cognitive load: intrinsic, extraneous, and germane.

Intrinsic cognitive load refers to the nature of the instructional materials, specifically the number of interacting elements that must be held in working memory for knowledge acquisition. For example, a novice learning the alphabet has low intrinsic

cognitive load as each letter may be learned in isolation. In contrast, learning to solve differential equations is a task with high intrinsic cognitive load due to the learners' need to simultaneously process the elements.

Extraneous cognitive load is imposed by the presentation of instructional materials, including the message design and instructional strategies (Kalyuga, Chandler, & Sweller, 1998). For example, requiring learners to search for and mentally integrate multiple sources of information, such as separate text and diagrams, places an increase in extraneous cognitive load.

Finally, germane cognitive load refers to the working memory resources needed for schema construction. Presentation strategies, such as providing learners with worked examples, allow for memory resources to be devoted to considering each problem state, increasing germane cognitive load, rather than considering a multitude of possible moves inherent in solving practice problems (Sweller, 2010). A discovery learning environment that has high intrinsic cognitive load, due to its complexity, and high extraneous cognitive load from lack of instructional guidance will result in little available working memory resources for schema construction (Kirschner, Sweller, & Clark, 2006; Sweller, 1999).

Cognitive load theory has primarily focused on strategies related to the message design of instructional materials in an attempt to reduce extraneous cognitive load and support schema acquisition. Currently, little research exists on how the incorporation of instructional guidance and supports, specifically generative strategies, may encourage the germane processes associated with schema construction and reduce the cognitive load experienced by novice learners.

Generative Strategies

Interacting in the guided discovery of principles, by means of responding to hints and working through examples, does not necessarily result in the deeper processing of content (Chi, 2009). Learners must be prompted to discover the *meaning* of a principle (Wittrock, 1979), by relating and integrating this new information with prior knowledge, a more constructive process (Chi, 2009). This constructive process may be facilitated through the incorporation of generative strategies with a guided discovery approach for principle learning.

Prompting learners to increase their depth of processing and meaning making of content can be accomplished through the use of integration or elaboration generative strategies as described by Jonassen (1988). Integration strategies facilitate the transformation of information into a more memorable form, such as paraphrasing, metaphors, and providing new examples. Integration strategies allow for the integration of information, restructuring, or refinement to existing schema, which is based on schema theory (Rummelhart & Ortony, 1977). Elaboration strategies prompt learners to add their own knowledge to new information, making it more meaningful. Elaboration strategies include generating mental or physical images, analogies, self-explaining, and implications. The following is a review of the generative strategies of paraphrasing and prediction with self-explanation that may be applicable to the learning of principles in addition to a guided discovery approach.

Paraphrasing. Paraphrasing and summarizing are terms that are often used interchangeably in research, though they reflect different cognitive processes. Summarizing involves the selection of important information, omitting details, collapsing

several related events into a single event, and generally reducing the overall length of text (Kintsch & van Dijk, 1978). This strategy allows for the organization of content in memory, but may not support the integration of prior knowledge with new information. In contrast, paraphrasing requires learners to use their own words and prior experiences to create novel sentences that reflect connections between prior knowledge and new information (Wittrock, 1989; Wittrock & Alesandrini, 1990). Paraphrasing may serve the purpose of increasing attentional processes to new information, as well as activating prior knowledge, leading to a deeper processing of material at a more meaningful level (Peper & Mayer, 1986). According to the generative model of learning (Wittrock, 1989), prompting learners to make meaningful connections between new information and prior knowledge by paraphrasing, memory for new information should be stronger than reading content without engaging in generative strategies.

A study by Glover, Plake, Roberts, Zimmer, and Palmere (1981) examined different note-taking strategies across two experiments. Undergraduate participants were randomly assigned to treatments where they read a text and simultaneously either (a) identified and documented key words, (b) identified verbatim, paraphrase, or conclusion statements related to the text, (c) paraphrased each paragraph, (d) created a conclusion statement for each paragraph, or (e) did not complete a strategy. Results of a free-recall test indicated that participants who actively paraphrased each paragraph remembered significantly more idea units than all other groups. During the scoring of free recalls, raters noted that participants did not fully complete the assigned strategy in the key words, paraphrase, and logical extensions groups. A second experiment sought to control for non-compliance among participants by establishing a participation requirement in

order for participants to receive credit and by evaluating participant materials prior to inclusion in the analysis. An analysis of free recall idea units in the second experiment revealed that the paraphrasing strategy again significantly improved recall, as did creating inference statements, when compared to the other strategy groups. Strategies that required learners to actively attend to text content and integrate with prior knowledge improved recall when compared with more passive reading and note taking strategies.

A similar comparison between note-taking strategies was examined by Bretzing and Kulhavy (1979) who explored depth of processing, review of notes, and immediate and delayed testing. High school participants summarized, paraphrased, copied verbatim sentences, completed a letter search, or simply read a passage on a contrived topic. Half of the participants in each note-taking group were allowed to review their notes prior to a posttest, whereas the remaining participants and the control group read an interpolated passage. Results of the immediate and delayed comprehension posttests favored participants in the summary and paraphrase groups, who performed significantly higher than the other groups. The review of notes for the paraphrase group did not significantly affect performance on the immediate test but did significantly improve performance on the delayed test. Results of this study confirm and extend the results from the Glover et al. (1981) study in that increasing the depth of processing during reading, facilitated by paraphrasing and summarizing, improves immediate and delayed retention. A concern of this particular study is the use of an interpolated passage for the no review and control groups. Reading an unrelated passage while other participants reviewed their notes may have affected content in working memory, which could explain the poorer performance by the control group. Participants in the summary and paraphrase no review groups,

however, did perform better on the immediate and delayed posttests than both the verbatim review and letter search review groups.

Improved effects for paraphrasing was further demonstrated by Wittrock and Alesandrini (1990), who compared the effects of paraphrasing or creating analogies to a control group who only read a lengthy passage. Undergraduate participants who paraphrased or created analogies for each paragraph performed significantly higher on a comprehension test than participants who read the passage without completing a strategy. The results of this study coincide with those reviewed previously that prompting learners to generate relations, through paraphrasing or analogies, facilitates learning. This was true across text components and between text and prior knowledge.

Paraphrasing has also been examined during acquisition of a procedure in a series of experiments by Glover, Timme, Deyloff, Rogers, and Dinell (1987). In one of the experiments in the series, undergraduate participants were read a set of 21 directions for assembling a distillation apparatus. After a step was read, half of the participants verbally paraphrased the step and then all participants were prompted to complete the step. After participants assembled the apparatus, they documented the steps of the procedure, which was scored by two raters for number of steps recalled and the correct order of steps. Participants in the paraphrase group recalled significantly more steps and more correctly documented the steps in order than the control group. The facilitative effects of paraphrasing each step of the procedure on recall of steps was replicated with high school students using a shorter set of directions and the same set of directions as the previously described experiment. A higher proportion of high school participants who paraphrased each step in a shorter set also correctly assembled the apparatus than those who did not

complete the generative strategy. This effect, however, was not observed for the lengthier set of directions. Results of this study extend the previous research on paraphrasing to procedure leaning and demonstrate this strategy to be an effective technique for the recall of a procedure for both high school and college students.

Summary and directions. Several conclusions can be drawn from the results of the studies on paraphrasing. Prompting learners to paraphrase text content and verbal directions facilitates acquisition and retention of instructional material. Learning is accomplished by encouraging a deeper processing of content, allowing for the integration between new and existing knowledge, as compared with strategies that facilitate a more shallow processing of content.

The studies reviewed have primarily concentrated on the learning of facts and concepts presented through text. What remains unknown in regards to paraphrasing is the effectiveness of this strategy in learning principles, as well as how this strategy may facilitate learning in a simulation environment. All of the studies reviewed have examined verbal learning, whether through reading a text or listening to a series of directions. The question remains as to how effective paraphrasing may be for the learning of principles in a simulation environment, which is primarily composed of non-verbal materials. Furthermore, the effect of paraphrasing on learners' perceived cognitive load, metacognitive judgments, and interest in content have been neglected in the existing body of research.

Prediction and self-explanation. An additional generative strategy that may be employed with a guided discovery approach of principle learning is prediction and self-explanation. Learners may first be prompted to make a prediction regarding the

relationship between concepts. After manipulating variables in a simulation environment, learners may self-explain the observed relationship and also rectify any discrepancies between an initial prediction and actual results.

Prediction during reading comprehension involves the activation of prior knowledge and experiences followed by confirming or disproving the prediction based on information presented in text (Collins, Brown, & Larkin, 1980; Palincsar & Brown, 1984). For example, learners may be asked to predict what may happen next in different parts of a story. In contrast to the generative strategy of paraphrasing after reading, prediction occurs prior to reading and allows for learners to anticipate the structure and content of upcoming information (Anderson & Pearson, 1984). Prediction encourages the integration of prior knowledge with information presented in the text and is consistent with schema theory (Rummelhart & Ortony, 1977). Prediction also facilitates comprehension monitoring since learners may, after reading text, realize an initial prediction was inaccurate and will then modify the prediction to more appropriately reflect the content in the text (Afflerbach & Walker, 1990).

The strategy of prediction may also have beneficial effects on learners' motivation. As observed by White and Gunstone (1992), the task of reasoning on possible assignment results by means of creating a prediction can be motivating for learners. For example, participants in a study by Lewis, Stern, and Linn (1993) reacted favorably to creating predictions using past experiences followed by investigating relationships with a simulation.

Prediction has found to be an effective strategy to improve reading comprehension (Freeman, 1982; Hansen, 1981), mathematical understanding and

reasoning (Kasmer & Kim, 2011), and learning from diagrams and animations (Byrne, Catrambone, & Stasko, 1999; Hegarty, Kriz, & Cate, 2003). In each of these studies, prompting learners to make predictions prior to or during instruction activated learners' prior knowledge and encouraged learners to make connections between existing knowledge and new information.

The facilitative effects of prediction may be improved by also requiring learners to self-explain. Self-explaining requires learners to go above and beyond the information presented (Chi & VanLehn, 1991) and prompts learners to actively construct understanding through the integration of new information with existing knowledge (Chi, De Leeuw, Chiu, & Lavancher, 1994). Self-explanations result in the generation of new content not presented in instructional materials whereas paraphrasing involves transforming provided information in the learner's own words (Hausmann & VanLehn, 2010). Self-explanations can vary from a shallow level, such as a description of an observation, or a deeper level containing a justification of why something happened (Okada & Simon, 1997). The latter requires learners to describe evidence supporting the explanation, involving a metacognitive component as learners must reflect on their comprehension to provide an appropriate explanation. An additional noted benefit of self-explanation is that it allows learners to devote cognitive resources relevant to germane cognitive load when dealing with instruction that has high intrinsic cognitive load (Sweller et al., 2011).

In a seminal study of self-explanations by Chi, Bassok, Lewis, Reimann, and Glaser (1989), college student participants studied a section of an introductory physics text with criterion testing and remedial instruction when necessary to ensure comparable

background knowledge. Participants then studied three worked examples of problem solving and verbalized their studying strategies, followed by a posttest requiring learners to solve physics problems. An examination of strategies verbalized while studying worked examples revealed that the participants who performed highest on the posttest elicited significantly more self-explanations related to physics concepts and principles presented in the text. The correlation between the number of self-explanations and performance on the posttest was high ($r = .81$). The high-performing students also produced significantly more comprehension monitoring statements than the low-performing students while studying worked examples. These monitoring statements by high-performing students were frequently followed by self-explanation statements. Results of this study suggest that prompting students to self-explain may aid in a more elaborate schema, enabling students to make connections between external information and integrate them with prior knowledge.

The effects of prediction with self-explanation when learning chess was examined in a study by de Bruin, Rikers, and Schmidt (2007). Undergraduate novice participants initially viewed a presentation on the basic rules of chess and then viewed a series of games played by a computer to induce the chess principles that underlie a specific endgame. During the viewing of the series of games, participants (a) observed the game, (b) predicted the next move the computer would make, or (c) predicted and self-explained the next move. Participants in the predict and self-explain group were also prompted to explain any discrepancies between their initial prediction and the actual move by the computer. Principle acquisition was assessed by having participants play five new examples of the endgame against the computer with as few moves as possible. An

analysis of interactions during the learning phase indicated the predict and self-explain group produced significantly more correct predictions and application of principles in predictions than the predict only and observe only condition. During the test phase, participants who predicted and self-explained applied principles significantly more often and made significantly fewer errors than the other two groups. Prediction and self-explanation facilitated principle understanding and transfer of principles learned to new examples as compared to prediction and observation or observation alone.

Summary and directions. Although several studies have demonstrated the advantages of prompting learners to create predictions, only one study has examined the combined effects of prediction and self-explanation on learning. This single study showed that principle acquisition and application were improved by prompting learners to predict and self-explain during a learning phase presented through an animated computer game. No additional studies could be identified in the research literature that examined the effects of prediction and self-explanation as compared with other strategies for principle learning. The effects of these strategies in an interactive simulation environment have also been neglected, as well as the combined effects of these strategies on participants' interest in the instructional content. Furthermore, it is unknown whether participants' subjective assessments of mental effort may differ when generative strategies are incorporated, an important factor related to cognitive load theory (Sweller et al., 2011).

Calibration

Calibration refers to the correspondence between a learner's perception of performance and a learner's actual performance (Keren, 1991). For example, calibration

may be assessed by having learners make a prediction regarding anticipated test performance and a postdiction of test performance (Bol, Hacker, O'Shea, & Allen, 2005). Prediction of anticipated test performance differs from the learning strategy of prediction described and reviewed in research previously. Prediction as a learning strategy involves the activation of prior knowledge and anticipation of to-be learned content. A learner's metacomprehension judgment in predicting test performance, however, involves a self-assessment of the knowledge he or she possesses, how thoroughly this content is understood, and the learner's ability to apply this knowledge during a test (Hacker, Bol, Horgan, & Rakow, 2000). Postdiction, on the other hand, refers to the monitoring judgment a learner makes in regards to how well he or she actually performed on a test.

Calibration accuracy, by means of prediction and postdiction of exam performance across a full semester, was examined in a study by Hacker, et al. (2008). Undergraduate participants either (a) reflected on explanations of calibration judgments, (b) were provided with extrinsic incentives to improve calibration accuracy, (c) reflected and were provided with incentives, or (d) merely predicted and postdicted exam performance. Lower-performing students significantly improved calibration accuracy when provided with incentives, whereas higher-performing students were consistent in calibration accuracy across the exams. It was revealed that higher-performing students relied most on their performance assessment in making predictions and postdictions. In addition to assessments, the attributional style constructs of internal studying behaviors and external social influences (e.g. instructor feedback on work, comparisons of performance with other students) were significant predictors of calibration judgments for lower-performing students. The results of this study suggest that learners may attribute

inaccurate calibration to insufficient studying practices, perhaps because learners may not spontaneously engage in strategies that encourage a deeper processing of content, as noted by Rigney (1976).

A later study by Bol, Hacker, Walck, and Nunnery (2012) examined the effects of guidelines at a half-way point during an exam review session in group or individual settings on calibration accuracy with high school students. Participants reflected on their understanding in either a group of five to six students or alone, with or without guidelines. The guidelines, which consisted of five questions, encouraged students to self-monitor learning and provided learners with self-regulatory feedback regarding how knowledge may be increased. Participants in a group setting with calibration guidelines exhibited the greatest calibration (prediction and postidiction) accuracy and achievement test scores. This increased accuracy in a group setting may be attributed to the encouraged reflection and the comparison of levels of understanding with peers, as well as the provision of guidelines shown to promote learning in other contexts.

Calibration accuracy in regards to prediction of test performance may be facilitated by increasing the learner's depth of processing, resulting in schema development. Prior research has suggested calibration accuracy is improved when learners engage in active processing of content such as filling in missing letters in words (Maki, Foley, Kajer, Thompson, & Willert, 1990), summarizing text content (Schommer & Surber, 1986; Thiede & Anderson, 2003), or self-questioning while reading (Davey & McBride, 1986).

Summary and directions. Results of the studies reviewed suggest that learners may attribute calibration judgments to internal studying strategies. Furthermore, engaging

in strategies that promote a deeper processing of content may improve calibration accuracy in addition to achievement. Research is currently lacking as to how alternate generative strategies, such as paraphrasing or prediction with self-explanations, may also affect calibration accuracy.

Purpose of Research

The purpose of this study was to extend previous research on generative strategies, specifically within an instructional simulation using a guided discovery approach for learning physics principles. The primary purpose was to determine whether generative strategies, paraphrasing and prediction with self-explanation, would improve learning of visually-presented principles in an interactive environment. Prior research has primarily examined paraphrasing and prediction with self-explanation on comprehension and application of verbal learning in non-interactive environments; this study investigated the effects of generative strategies in learning visually-presented material on recall, application, evaluation, and transfer problems. This study also examined the relative effects of generative strategies on perceived cognitive load, calibration accuracy, and interest, factors that have not been investigated in prior research.

Consistent with the generative model of comprehension proposed by Wittrock (Wittrock, 1974, 1989), it was hypothesized that generative strategies (paraphrasing, prediction + self-explanation) would improve guided discovery learning as compared with a control group who received only guided discovery during learning. Additionally, since calibration accuracy has been shown to improve with a greater depth of processing (e.g. Thiede & Anderson, 2003; Davey & McBride, 1986), a second hypothesis proposed

that learners engaging in generative strategies would exhibit greater calibration (prediction and postdiction) accuracy as compared with the control group.

Three exploratory research questions were also examined. The first examined the relationship between quality of paraphrases and self-explanations with test performance. Specifically, the relationship between the thoroughness of generative content and test performance was examined, as well as the relationship between the depth of self-explanations and test performance. The second research question investigated how generative strategies affect perceived cognitive load. It was anticipated that participants engaging in generative strategies would report higher levels of mental effort than participants in the control group due to the increased depth of processing encouraged by the strategies. The third research question explored how reported interest in the instructional content might differ among the three groups.

CHAPTER II

METHOD

Design

This study employed a true-experimental design. The independent variable was the generative strategy (none, paraphrasing, prediction+self-explanation). Dependent variables were performance on an achievement test, calibration accuracy, perceived cognitive load, generative content quality, and interest towards materials.

Participants

Eighty-five undergraduate participants (57 women and 28 men) were recruited from students enrolled in education courses at a southeastern university. The average age of participants was 20.55 years (*SD* = 2.46). To minimize the effect of prior knowledge in the content area, participants who cited prior high school or college level course-work in physics were excluded from participation in the study. Participation was voluntary, though extra credit and a small monetary reimbursement was offered in exchange for participation. A delayed ruse was employed to ensure participant effort during this study. Participants were informed that a minimum score of 70% was required to receive the extra credit incentive. After all data had been collected, a debrief notice was immediately emailed to all participants explaining the ruse and informing them that all have received extra credit regardless of performance.

Materials

Instructional materials were first pilot tested before being implemented in this study. A class of 17 undergraduate students completed the control group materials and responded to test items. This pilot testing allowed for the evaluation of test reliability,

individual test item discrimination indices, and to provide an approximation of instructional time.

Instruction. Instructional materials consisted of five sequential, concrete assignments for completion within the instructional simulation (see Appendix A). Each assignment began with a statement for the purpose of the task (e.g., Discover the relationship between launch angle and projectile distance). Participants then manipulated individual variables within the simulation (see Appendix B) such as projectile mass, launch height, launch angle, and initial velocity, documented effects on flight time and distance, and determined the relationship between the concepts.

Measures

Achievement test. An achievement test was administered that consisted of 34 multiple-choice (n = 22) and short-answer (n = 12) items measuring recall (n = 10), application (n = 10), evaluation (n = 5), and near transfer (n = 9) of principles learned. A test blueprint was developed (see Appendix C) and the test items (see Appendix D) were reviewed by three experts to establish content validity. The achievement test was piloted with a class of 17 students who completed instructional materials for the control group and resulted in an internal consistency reliability of .88 as calculated with Kuder-Richardson Formula 20 (KR-20). Internal consistency reliability of the test in this study with all 85 participants was .87 using the KR-20.

Calibration. Calibration accuracy was assessed through prediction accuracy and postdiction accuracy of test scores. Participants responded to a test prediction question (What raw score do you anticipate receiving on this test?) and a test postdiction question (Now that you have taken the test, what score do you think you will receive?).The

absolute difference was calculated between participants' prediction and actual test performance, as well as the absolute difference between participants' postdiction and actual test performance.

Generative content. Two raters independently evaluated participants' quality of paraphrases and self-explanations in the two generative strategy groups. Any discrepancies were resolved through discussion and a single score was determined. First, the raters examined the written responses by each participant and compared their responses to a rubric (see Appendix E) to determine the number of idea units present in paraphrases and self-explanations. Idea units ranged from a possible one to three idea units per principle with a total possible idea unit count of nine idea units for all five principles. A quality score was calculated for each participant based on the number of idea units present out of the total possible idea units.

Second, raters examined the quality of the self-explanations, which contained comparisons between initial predictions and actual results, as well as potential reasons for actual results. A score of 0 was given to any explanation that either did not contain an explanation for the results of the assignment or for an explanation that was descriptive (e.g. "I think the distance didn't differ because the launch angle stayed the same"). A score of 1 was given when an explanation provided a possible cause to justify the results of the assignment (e.g. "This could be explained with gravity and how it pulls any object, regardless of mass, down at the same rate").

Cognitive load. Subjective assessments of task demands was measured with the NASA-TLX questionnaire originally developed by Hart and Staveland (1988) and incorporated modifications as employed in a study by Gerjets, Scheiter, and Catrambone

(2006).The questionnaire (see Appendix F) consisted of four subscales: mental effort, mental demand, performance, and frustration. Items for mental effort, mental demand, and frustration were rated on a scale ranging from 0 (very low) to 100 (very high). The performance items were rated on a scale ranging from 0 (good) to 100 (poor). Participants responded to the single mental effort question (How hard did you have to work in your attempt to understand the contents in the instruction?) after completing each of the five assignments. Administering a subjective rating scale multiple times during a learning task may provide insight into variations of experienced mental effort over time (Antonenko, Paas, Grabner, & van Gog, 2010).

At the end of instruction, participants responded to questions from the remaining three subscales, mental demand ($n = 2$), performance ($n = 2$), and frustration ($n = 1$). It has been proposed by Gerjets et al. (2006) that the subscales of mental demand and effort translate to the dimensions described by cognitive load theory (Sweller et al., 2011). The mental demand subscale describes the cognitive activity needed to understand the learning task, which may relate to the complexity of instructional materials, or intrinsic cognitive load. The effort subscale describes how hard the learner must work to understand the content, or the cognitive activities related to germane processes. Test-test reliability correlation of the original questionnaire was .83 (Hart & Staveland, 1988). Reliability of the mental effort question administered five times in this study was $\alpha = .89$, and the reliability of the remaining three subscales ($n = 5$) together was $\alpha = .80$.

Interest. Participants' interest in the content presented in the instructional materials was measured with an adaptation of the Perceived Interest Questionnaire (PIQ) (Schraw, Bruning, & Svoboda, 1995). The questionnaire (see Appendix G) consisted of

10 items rated on a 5-point Likert-type scale (ranging from 1 = "strongly disagree", 5 = "strongly agree"). A total interest score was determined by calculating the average across all item responses. Internal consistency of the interest questionnaire in the original study was α = .91. Reliability of the questionnaire in this study was α = .82.

Procedure

The experiment was conducted during scheduled times in a computer lab. In an effort to counteract diffusion of treatment effects, participants at designated scheduled times were randomly assigned to one of three treatments: a control group (*n* = 28), a paraphrase group (*n* = 28), or a prediction+self-explanation group (*n* = 29). An example assignment illustrating the differences between treatment groups is presented in Table 1.

Table 1

Example assignment by treatment group.

Control	Paraphrase group	Prediction+self-explanation group
Discover how launch angle affects projectile distance and flight time.	Discover how launch angle affects projectile distance and flight time.	Discover how launch angle affects projectile distance and flight time. How do you predict launch angle will affect the distance traveled?
Set angle: Distance: Time: 15° 30° 45° 60° 75° 90°	Set angle: Distance: Time: 15° 30° 45° 60° 75° 90°	Set angle: Distance: Time: 15° 30° 45° 60° 75° 90°
	In your own words, explain the relationship between angle and distance and flight time. Explain the results of the experiment by relating it to your own paraphrase.	How do the results of this experiment compare with your initial prediction? Explain why the results confirmed or disproved your initial prediction.

Participants in all groups first received an example assignment (see Appendix H). In addition to the example assignment, participants in the generative strategies groups also received an example for each of their assigned strategies. Those in the paraphrase group completed an assignment and were prompted to paraphrase what they observed and the principle learned. Those in the prediction+self-explanation group first made a prediction as to the relationship between the variables, completed the assignment, and then self-explained differences between their prediction and observed results and provided a possible justification for results.

Participants individually accessed the assignments and instructional simulation through a website on a lab computer. They worked through the instructional materials and responded to the mental effort question after completing each of the five assignments. They then responded to the remaining items on the cognitive load questionnaire, completed the interest questionnaire, made a prediction of anticipated test performance, and completed the test. Finally, they made a postdiction on test performance. Participants worked through the instructional materials at their own pace and instructional start and end time were recorded. Prior research (e.g. Wittrock & Alesandrini, 1990) suggests that differences in instructional times between groups with and without generative strategies are not statistically significant and do not significantly contribute to learning differences. Participants in the control group were allowed the opportunity to immediately review instructional materials to equate instructional time. All data were collected online.

A one-way analysis of variance (ANOVA) was conducted to first determine if there were significant differences in time-on-task between groups. The results of the

analysis failed to indicate a statistically significant difference between the control group ($M = 32.58$, $SD = 3.98$), paraphrase group ($M = 37.20$, $SD = 10.10$), or the prediction+self-explanation group ($M = 37.22$, $SD = 10.26$), $F(2,82) = 2.70$, $p > .05$, $\eta^2 = .06$. Results of this analysis rule out time as a confounding variable.

CHAPTER III

RESULTS

This chapter presents the results of the analyses used to evaluate the effects of generative strategies on achievement and calibration accuracy. This is followed by a presentation of the results related to generative content created by the two treatment groups. Last, results regarding dimensions of cognitive load and reported interest between the three groups is presented.

Analysis of Test Performance – Hypothesis 1

A one-way multivariate analysis of variance (MANOVA) was conducted to evaluate the differences between the three groups on the four dependent variables of test item levél. Table 2 presents means and standard deviations on recall, application, evaluation, and transfer test items for the three groups.

Table 2

Means and standard deviations of test scores by item type

		Test item level							
		Recall		Application		Evaluation		Transfer	
Group	*n*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Control	28	4.79	1.87	5.18	2.16	2.29	1.41	3.36	1.97
Paraphrase	28	6.93	2.12	6.39	1.87	3.32	1.42	5.04	2.04
Prediction+self-explanation	29	7.10	1.76	6.34	1.88	3.59	1.21	4.93	2.23

Note: Scores could range from 0 to 10 for recall items, 0 to 10 for application items, 0 to 5 for evaluation items, and 0 to 9 for transfer items.

The results of the analysis revealed a significant difference in test performance, Wilk's $\Lambda = .74$, $F(8,158) = 3.18$, $p = .002$, multivariate $\eta^2 = .14$. Analyses of variance

(ANOVA) on each item type were conducted as follow-up tests to the MANOVA. Significant differences were found between groups on recall items, $F(2,82) = 12.70$, $p<.001$, $\eta^2 = .24$, application items, $F(2,82) = 3.42$, $p=.038$, $\eta^2 = .08$, evaluation items, $F(2,82) = 7.36$, $p=.001$, $\eta^2 = .15$, and transfer items, $F(2,82) = 5.71$, $p=.005$, $\eta^2 = .12$.

Follow-up comparisons using the Tukey HSD procedure showed that both of the generative strategy groups significantly exceeded the control group on recall items, evaluation items, and transfer items, but not on the application items. The generative strategy groups did not differ significantly from each other on any of the four scales.

Calibration Accuracy – Hypothesis 2

A MANOVA was conducted to evaluate the effects of treatment group on calibration accuracy. Calibration accuracy was determined by the absolute difference between prediction and actual test performance and the absolute difference in postdiction and actual test performance. Since calibration is determined by differences between prediction and postdiction scores and actual test scores, higher scores for calibration indicate less accuracy. Table 3 illustrates the means and standard deviations for test prediction accuracy and test postdiction accuracy for the three treatment groups.

Table 3

Calibration accuracy by treatment group

	Calibration									
	Actual test score		Prediction score		Prediction accuracy		Postdiction score		Postdiction accuracy	
Group	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Control	45.90	17.67	82.04	8.59	36.13	19.74	63.86	16.58	20.26	15.92
Paraphrase	63.76	19.26	83.04	9.08	21.48	14.68	70.11	15.45	13.78	10.68
Prediction+ self-explanation	64.60	16.31	79.38	7.98	17.16	12.74	72.83	11.28	12.56	9.85

The MANOVA indicated statistically significant differences between groups, Wilk's $\Lambda = .79$, $F(4,162) = 5.21$, $p = .001$, multivariate $\eta^2 = .114$. Analyses of variance (ANOVA) on the dependent variables were conducted as follow-up tests to the MANOVA. The ANOVA for prediction accuracy was significant, $F(2,82) = 10.99$, $p<.001$, $\eta^2 = .211$ as was postdiction accuracy, $F(2,82) = 3.14$, $p = .049$, $\eta^2 = .071$.

Tukey multiple comparisons indicated that the generative strategy groups were significantly more accurate in their prediction scores than the control group. There were no significant differences between the two generative strategy conditions. A near-significant trend ($p = .056$) was found for postdiction accuracy between the control and prediction+self-explanation group. No other differences were statistically significant.

Content Quality – Research Question 1

Across the five principles, the paraphrase group scored on average 6.35 (SD = 1.28) total idea units and the prediction+self-explanation group scored on average 5.72 (SD = 1.10) total idea units. The group difference in the number of idea units counted was statistically significant, $F(1,56) = 4.014$, $p = .05$. A Pearson correlation was computed between the number of idea units identified by participants in the two treatment groups and their achievement test scores. A positive and significant correlation was found, ($r(57) = .59$, $p<.001$), indicating a substantial relationship between the thoroughness of paraphrases and self-explanations and test performance.

Additionally, an examination of the quality of self-explanations revealed that those in the self-explanation group provided a potential explanation for the cause of the assignment results on average 2.28 (SD =1.46) out of five times. A Pearson correlation

was computed between the explanation quality of participants in the self-explanation group and their achievement test scores. A positive and significant correlation was also found, ($r(29) = .78, p<.001$), indicating a marked relationship between the depth of quality in self-explanations and test scores.

Considering the strong correlation between self-explanation quality and total test scores, additional exploratory analyses within the prediction+self-explanation group were performed. In a similar method to de Bruin et al.(2007), a median split was computed based on the number of self-explanations containing an assignment justification. The high-explainers (n = 14) consisted of all self-explainers with three or more justification self-explanations. The high-explainers provided an average of 3.57 (SD = .65) justifications. The low-explainers (n = 15) consisted of all self-explainers with less than three justification self-explanations. The low-explainers documented an average of 1.07 (SD = .80) justifications.

To evaluate potential differences in test performance between low- and high-explainers, a one-way multivariate analysis of variance (MANOVA) was conducted. Table 4 presents means and standard deviations on recall, application, evaluation, and transfer test items between the two groups of explainers.

Table 4.

Means and standard deviations of test item types for low- and high-explainers

		Test item level							
		Recall		Application		Evaluation		Transfer	
Group	*n*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Low-explainers	15	6.40	1.76	5.67	1.99	3.00	2.19	4.04	1.25
High-explainers	14	7.86	1.46	7.07	1.49	5.86	1.96	4.21	.80

Note: Scores could range from 0 to 10 for recall items, 0 to 10 for application items, 0 to 5 for evaluation items, and 0 to 9 for transfer items.

The results of the analysis revealed a significant difference in test performance, Wilk's $\Lambda = .64$, $F(4,24) = 3.68$, $p = .025$, multivariate $\eta^2 = .36$. Analyses of variance (ANOVA) on each item type were conducted to investigate differences by scale. Significant differences were found between groups on recall items, $F(1,28) = 5.82$, $p = .023$, $\eta^2 = .18$, application items, $F(1,28) = 4.58$, $p=.042$, $\eta^2 = .15$, evaluation items, $F(1,28) = 9.50$, $p = .005$, $\eta^2 = .26$, and transfer items, $F(1,28) = 5.37$, $p = .028$, $\eta^2 = .17$. The high-explainers performed significantly higher on all item types than the low-explainers.

In light of the differences in test performance, calibration accuracy between the low- and high-explainers was examined with a one-way MANOVA. Table 5 shows the means and standard deviations for test prediction accuracy and test postdiction accuracy for the two groups of self-explainers.

Table 5

Calibration accuracy by low- and high-explainers

	Calibration									
	Actual test score		Prediction score		Prediction accuracy		Postdiction score		Postdiction accuracy	
Group	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Low-explainers	56.27	15.68	77.53	9.36	22.98	15.20	67.67	11.68	17.74	10.20
High-explainers	73.53	11.93	81.36	5.88	10.94	7.79	78.36	7.95	7.01	5.74

The MANOVA indicated statistically significant differences between groups, Wilk's $\Lambda = .69$, $F(2,26) = 5.92$, $p = .008$, multivariate $\eta^2 = .313$. Analyses of variance (ANOVA) on the dependent variables were conducted as follow-up tests. The ANOVA on prediction accuracy was significant, $F(1,28) = 8.11$, $p = .008$, $\eta^2 = .23$, favoring the high-explainers in prediction accuracy over the low-explainers. The ANOVA on postdiction accuracy was also significant, $F(1,28) = 11.93$, $p = .002$, $\eta^2 = .31$, again favoring the high-explainers. A parallel analysis on paraphrases was not conducted due to the lack of variance in overall paraphrase quality.

Cognitive Load – Research Question 2

A 3 (treatment groups) X 5 (trials) repeated-measures ANOVA was conducted to determine whether there were significant differences on the mental effort measure across trials representing different physics principles. Table 6 shows the means and standard deviations of the average mental effort scores for each treatment group.

Table 6

Mental effort scores during instruction

		Mental Effort	
Group	*n*	*M*	*SD*
Control	28	25.75	22.27
Paraphrase	28	30.88	20.33
Prediction+self-explanation	29	40.38	23.91

Note: Mental effort score could range from 0 = very low to 100 = very high.

Results from the repeated-measures ANOVA indicated a significant main effect for trials $F(4,79) = 7.31$, $p<.001$, multivariate $\eta^2 = .27$ and a significant main effect for treatment group, $F(2,82) = 4.75$, $p = .011$, multivariate $\eta^2 = .10$. There was also a

significant interaction effect, illustrated in Figure 1, between mental effort ratings by trial and treatment groups, $F(8,158) = 3.45, p = .001$, multivariate $\eta^2 = .15$.

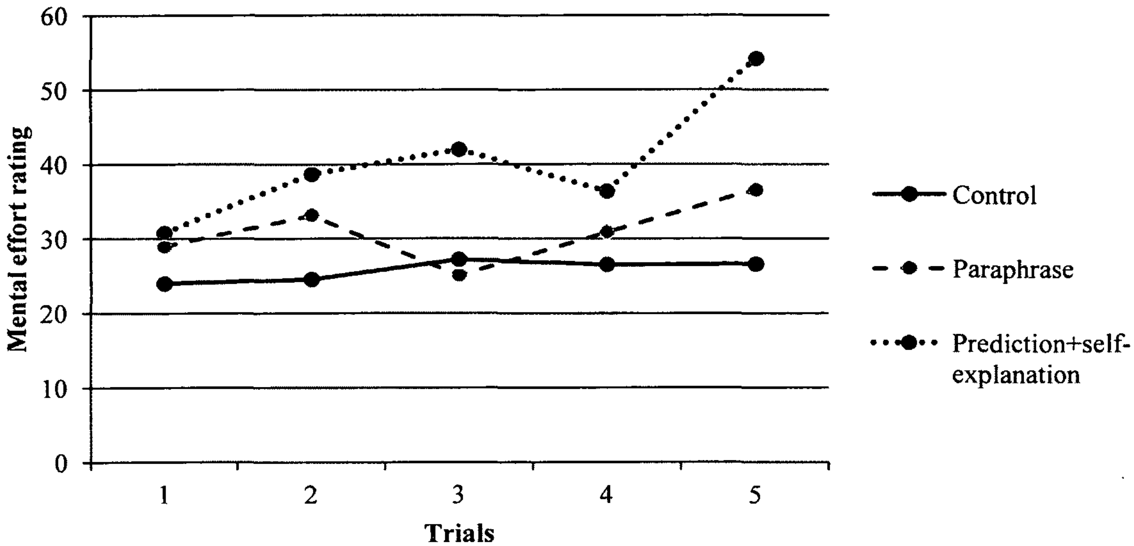


Figure 1. Mean mental effort ratings by each of the five principles.

Follow-up analyses were conducted through Tukey's multiple comparisons on the significant interaction effect. Results indicated that the prediction+self-explanation group reported significantly higher mental effort ratings than the control group for the second, third, and fifth principle. Additionally, the prediction+self-explanation group reported significantly higher mental effort ratings than the paraphrase group on the third and fifth principles. No other differences were statistically significant.

As described previously, the items for the subscales of demand, performance, and frustration from the cognitive load questionnaire were administered once at the end of instruction. These items were analyzed with a MANOVA. Means and standard deviations

for each of the subscales, demand, performance, and effort, by treatment group are shown in Table 7.

Table 7

Means and standard deviations for mental demand, performance, and frustration.

	Subscale					
	Demand		Performance		Frustration	
Group	*M*	*SD*	*M*	*SD*	*M*	*SD*
Control	35.27	18.68	26.29	23.89	15.39	20.76
Paraphrase	43.91	23.30	19.13	14.25	29.93	25.03
Prediction+self-explanation	48.40	20.86	32.41	20.53	34.45	25.25

Note: Demand and frustration scores could range from 0=very low to 100=very high. Performance scores could range from 0=good to 100=poor.

Results of the MANOVA revealed a statistically significant difference in the questionnaire responses, Wilk's $\Lambda = .176$, $F(6,160) = 3.16$, $p = .006$, multivariate $\eta^2 =.11$. Analyses of variances (ANOVA) were conducted as follow-up tests to the MANOVA. Significant differences were found on the performance subscale, $F(2,82) = 3.16$, $p = .048$, $\eta^2 = .07$ and the frustration subscale, $F(2,82) = 4.96$, $p = .009$, $\eta^2 = .11$. Differences on the demand subscale approached significance, $F(2,82) = 2.86$, $p = .06$, $\eta^2 = .07$.

Tukey's multiple comparisons were conducted as follow-up tests to the univariate ANOVAs for the performance subscale and the frustration subscale. The prediction+self-explanation group was significantly less confident in their performance than the paraphrase group. No other comparisons on the performance subscale were significant. The prediction+self-explanation group reported significantly higher levels of frustration than the control group. No other comparisons reached statistical significance.

Interest – Research Question 3

Differences in reported interest towards instructional material were analyzed through a one-way ANOVA. The independent variable included three conditions: control ($M = 3.1$, $SD = 0.6$), paraphrase ($M = 3.2$, $SD = 0.6$) and predict+self-explain ($M = 3.2$, $SD = 0.6$) and the dependant variable was the mean of participant responses on the interest questionnaire. The results failed to show a statistical difference in interest between groups, $F(2,82) = .091$, $p > .05$, $\eta^2 = .002$.

An examination of individual items, displayed in Table 8, reveals differences between groups worthy of note. The prediction+self-explanation group reported the highest level of agreement to the questionnaire item asking whether participants were caught-up in the instruction without trying. Also, participants in the paraphrase group reported the highest level of agreement to a question on a desire to learn additional information in the content area. Responses to a questionnaire item regarding a desire to complete the instruction again indicate the highest level of agreement by the control group. This may be due to students in the control group reporting less frustration with the task on the cognitive load questionnaire.

Table 8

Mean interest questionnaire responses.

	Treatment Groups					
	Control		Paraphrase		Prediction+self-explanation	
Questionnaire items	*M*	*SD*	*M*	*SD*	*M*	*SD*
I would complete this instruction again if I had the chance.	3.54	1.07	3.00	0.94	2.96	0.88
I got caught-up in the instruction without trying to.	3.07	1.02	3.04	0.88	3.36	1.06
I would like to learn more about this topic in the future.	3.14	0.88	3.38	0.94	2.93	1.13

CHAPTER IV

DISCUSSION AND CONCLUSIONS

Significant Findings

The purpose of this research was to examine the effects of generative strategies on principle learning, calibration accuracy, and dimensions of cognitive load in a simulation environment. Participants completed five assignments that employed a guided discovery approach for learning physics principles related to projectile motion. Participants either (a) completed assignments without generative strategies, (b) completed assignments and paraphrased, or (c) created predictions, completed assignments, and self-explained. In this chapter, the results are explained and their implications for future practice and research are discussed.

Test Performance

Results of this study provided support for the hypothesis that prompting learners to engage in generative strategies would improve learning of physics principles presented within the simulation. Participants who paraphrased or predicted and self-explained assignment results exhibited better performance on all measures of the achievement test as compared with participants who merely conducted the prescribed simulation manipulations. This improvement in learning is attributed to a greater depth of processing of content as described by Wittrock's (1974, 1989) generative model of learning regarding the importance of prompting learners to make connections between new information and prior knowledge or past experiences.

Additionally, results of this study are consistent with the argument by Chi (2009) that knowledge construction is superior to merely interacting with instructional materials,

in this case, manipulating variables to complete assignments in a guided discovery approach. These results expand on the benefits of paraphrasing found in past print-based research (e.g., Glover et al., 1981; Wittrock & Alesandrini, 1990) and prediction with self-explanations from a computer game (e.g., de Bruin et al., 2007) to principle learning in a simulation environment.

The improvement in learning due to engaging in generative strategies may be best understood through the examination of participants' generative content. A strong correlation ($r = .59$) was found between the thoroughness of paraphrases and self-explanations, as assessed through idea unit count, and test performance. Although there were a significantly greater number of idea units in the paraphrases than the self-explanations, test performance did not differ between these two groups. This lack of difference is possibly due to the nature of the tasks inherent in the strategies. Paraphrasing led to more thorough generative content, as reflected in the greater amount of idea units. The self-explanation strategy may have led to more concise generative content, but contained a greater depth than paraphrases due to the need to provide an explanation. Furthermore, the lack of difference in test performance between the two groups may be attributed to the added benefit of having learners make a prediction of the potential relationship in addition to self-explaining results of their assignments. Creating a prediction may have served to activate prior knowledge and focus attention on the integration of upcoming information, consistent with past research on the benefits of prediction during reading (Afflerbach, 1990).

The quality of self-explanations was also strongly correlated with test performance ($r = .78$). Self-explanations varied in quality as either descriptive (e.g.

"When something travels further then the flight time will obviously be longer,") or justification (e.g. "When you put a lot more power into the projectile, by increasing the velocity, there is nothing but the air that stops the projectile till it hits the floor, resulting in an increase in distance, height, and time,"). Less than half (45.6%) of all self-explanations were categorized as containing inferences of new information, characteristic of deeper self-explanations as compared with descriptive self-explanations (Okada & Simon, 1997). Although participants apparently struggled to create the deeper self-explanations, as evident in the low quantity of justification self-explanations present, performance was the highest for this group for three out of four measures on the achievement test. Those participants who did attempt to create a greater number of deeper self-explanations, however, performed significantly higher on all measures of the test compared to participants whose self-explanations tended to be more descriptive (e.g., summarizing). This finding coincides with previous studies in which higher-quality self-explanations were associated with improved learning (Chi et al., 1989; Okada & Simon, 1997; Renkl, 1997).

Calibration Accuracy

The hypothesis that engaging in generative strategies would improve calibration (prediction and postdiction) accuracy, a measure of self-regulated learning, was partially supported. Learners in the generative strategy groups were significantly more accurate in predicting their performance than participants who did not produce generative content. Since generative strategies resulted in enhanced performance, students in this condition may be more accurate in predicting their depth of understanding and their ability to apply this information on the achievement test. This finding is consistent with past research by

Schommer and Surber (1986) and Maki et al. (1990) where calibration (prediction) accuracy was improved by increasing the depth at which content is processed. The creation of generative content in this study may have provided the feedback learners needed during the performance phase of Zimmerman's (2000) model of self-regulated learning to aid in an assessment of understanding. In contrast, the control group received no means to self-assess their understanding, which may have affected their ability to accurately predict test performance. This lack of feedback experienced by the control group is similar to a finding by Bol et al., (2012) where participants that received guidelines prompting for reflection of understanding were more accurate in predicting test performance than those who did not receive guidelines.

Although participants in the generative strategy groups were more accurate in predicting test scores, postdiction accuracy did not differ significantly between groups in this study. This finding reflects those of other researchers (e.g., Bol et al., 2005) where postdiction accuracy was relatively stable across achievement level. As noted by Bol et al. (2005), postdictions of performance may be more accurate than predictions of performance due to participants' experience with the specific test items presented. Participants may use feedback provided by the test itself to produce more accurate postdiction of performance. Creating predictions of test performance, on the other hand, involves a variety of factors, including not knowing exactly what content will be tested and how this content will be assessed. However, postdiction accuracy directionally favored participants in the prediction+self-explanation group. Furthermore, both generative strategy groups were more accurate in their postdiction when compared to the control participants. In fact, the differences in postdiction accuracy between the

prediction+self-explanation group and the control group approached statistical significance ($p = .056$). The effect size was $d = .58$, indicating a moderate difference between groups.

The overall improved calibration accuracy by participants in the prediction+self-explanation group may be partially explained by the prompt to create a prediction for the anticipated relationship between variables. More than half (56%) of all predictions created by those in the prediction+self-explanation group were incorrect prior to the guided discovery of the principle represented in the activity. The prediction strategy served to activate prior knowledge by prompting learners to create an intuitive guess. As learners then encountered information from the simulation assignments that contradicted their prediction, this new information was integrated with prior knowledge and their existing mental model was adjusted accordingly. Participants were able to monitor their understanding of the physics principles as initial misconceptions were corrected, both through conducting the assignments and by comparing their predictions to the assignment results in self-explanations. For example, in the discovery that mass has no effect on the distance or height a projectile travels, one participant commented, "I think my way of thinking was wrong. I thought the power which sets the projectile up would be the same." Such comprehension monitoring statements may have facilitated calibration accuracy.

The difference in quality of self-explanations sheds further light on calibration accuracy with this particular group of learners. Participants who produced a greater number of justification self-explanations were significantly more accurate in both predicting and postdicting test performance than those who provided primarily descriptive self-explanations. This finding lends additional support to the notion that

deeper, or justification self-explanations require the learner to assess his or her understanding in order to provide a sound explanation for why something occurred (Okada & Simon, 1997). The task of creating a *why* explanation may have further facilitated learners' ability to accurately assess their understanding as measured through calibration.

Cognitive Load

Aspects of cognitive load, specifically effort, demand, performance, and frustration, were affected by the incorporation of generative strategies in this study. First, mental effort was significantly higher for participants in the prediction+self-explanation group as compared with the other two groups. The increased reported mental effort by participants creating predictions and self-explanations was not surprising. It was expected that perceived mental effort would be increased due to the attention directed to germane resources, specifically a greater depth of processing as related to schema construction (Sweller et al., 2011). This study is an important contribution to the body of research regarding cognitive load theory (Sweller et al., 2011), as it is one of the first to measure the effects of reported mental effort in relationship to strategies for increasing germane resources. Prior research guided by cognitive load theory (for a review, see Sweller et al., 2011) that measured reported mental effort has purported that an increase in mental effort is related to intrinsic cognitive load, leading to a detriment in learning. In this study, an increase in mental effort was more likely related to germane resources and actually benefited learning.

In contrast to those participants who predicted and self-explained, participants who paraphrased did not report significantly higher levels of mental effort during

learning as compared with the control group. These participants, however, exhibited equivalent scores on the achievement test as those in the prediction and self-explanation group. Why were mental effort ratings significantly higher for the prediction+self-explanation group but not for participants who paraphrased? A possible explanation is the perception of mental effort was increased due to the need for the prediction+self-explanation participants to complete two strategies as compared to the single task required in the paraphrase group. The variations in reported mental effort between groups may also be explained by the Amount of Invested Mental Effort (AIME) as described by Salomon (1981).

AIME (Salomon, 1981) refers to the intentional information processing that leads to learning, specifically the mental elaboration a learner engages in when presented with instructional materials. Salomon (1981) proposed that when mental effort is increased, learning will also increase. Furthermore, mental effort may be related to the perceived demand characteristics of the instructional materials. That is, when a medium (e.g., television) or instructional task is familiar, it is perceived as having lower demand characteristics, requiring less mental effort and therefore less is expended, often resulting in a decrease in learning (Salomon, 1981). The relationship between mental effort and demand characteristics was originally developed to explain differences in mental effort and achievement between various mediums such as television and print (e.g., Salomon, 1984). Since its initial development, AIME (Salomon, 1981) has also been used to explain differences in learning through a single medium when varying task orientation strategies were employed that may have influenced demand characteristics (Kunkel & Kovaric, 1983; Salomon & Leigh, 1984). It is therefore possible that AIME (Salomon,

1981) may explain the differences in mental effort related to the demand characteristics of instructional strategies employed. In this study, the correlation between ratings of demand and the average mental effort ratings was $r = .73$, lending support for the application of AIME to explain results.

While reported demand did not statistically differ between groups, the assessments of demand mirrored the order of increasing mental effort between the control group, paraphrase group, and prediction+self-explanation group. The levels of demand perhaps relates to the familiarity of the task. For example, the control group may have perceived the task of merely manipulating variables and conducting assignments with the simulation as fairly familiar and invested less mental effort. In contrast, the prediction+self-explanation group, whose ratings of demand were higher, may have had less experience creating self-explanations and expended a greater amount of mental effort during learning. Paraphrasing, in contrast, could have been a relatively familiar strategy to learners. These participants therefore expended an amount of mental effort greater than the control group but less than the prediction+self-explanation group.

There were significant differences on the performance subscale in this study. According to Salomon (1981), AIME is influenced not only by the perceived demand characteristics of the instructional materials, but also the learner's perceived self-efficacy. Perceived self-efficacy is described as a learner's belief, or confidence, in their ability to perform activities and aids in determining how much effort is exerted to accomplish these activities (Bandura, 1977). Whereas Bandura (1977) described a positive linear relationship between self-efficacy and effort, Salomon (1981) proposed a curvilinear relationship, where an increase in perceived self-efficacy is related to an increase in the

effort invested up to a certain point. Beyond this point, learners may perceive the instructional materials or activities as easy and then invest less effort as they are highly confident in their abilities. This curvilinear relationship proposed by Salomon (1981) may partially help to explain the significant differences between the control group and the prediction+self-explanation group on subjective measures of performance. Findings from this study indicated that the control group reported the lowest perceived demand and reported equivalent confidence in their performance to the paraphrase group. In contrast, the prediction+self-explanation participants reported the highest demand, as well as mental effort, but were the least confident on the performance subscale. While the research reported by Salomon (1981) measured perceived self-efficacy prior to instruction as related to perceptions of learning from various forms of media (e.g., television, print), this study measured perceptions of performance after instruction within a single medium. The proposed relationship by Salomon (1981), however, where a learner may feel less confidence in meeting the demands of a task that appears to require more effort was somewhat observed in this study. A moderate relationship between average mental effort ratings and performance ratings was observed in this study (r = .47), supporting Salomon's (1981) proposed relationship.

The final aspect of the cognitive load measure that resulted in differences between groups was the frustration subscale. Participants in the prediction+self-explanation reported significantly higher levels of frustration than those in the control group. Although other differences were not significant, again, similar to the demand subscale, the order mirrored that of increasing mental effort between the control group, paraphrase group, and prediction+self-explanation group. A potential explanation for the increased

levels of frustration reported by the prediction+self-explanation group could be due to the difficulties learners experienced creating self-explanations in an unfamiliar domain. Despite the prompt for learners to explain the differences between their prediction and assignment results and *why* a difference did or did not exist, learners were often unable to produce an appropriate explanation. For example, one participant commented, "I am not sure why this is so," whereas another wrote, "I cannot even begin to explain what just happened my mind is blown. I am sorry." Additionally, learners were not provided feedback regarding the accuracy of their self-explanations. This feeling of unknowing may have influenced their levels of frustration, since they were prompted to consider the results of the assignment at such a deep level without having prior knowledge to refer to in order to substantiate assignment results. A level of uncertainty regarding self-explanations may also help explain the lower confidence levels, indicated in the prediction+self-explanation group results on the performance subscale.

Interest

Results of this study indicated no difference in reported interest towards the instructional materials between groups. That is somewhat surprising in regards to the prediction+self-explanation group. As noted by prior researchers (e.g. Gunstone & White, 1981; Lewis et al., 1993) reasoning on possible relationships between variables and then confirming or disproving predictions through experimentation may be motivating for learners. This finding was not observed in the present study and may be perhaps due to the increased levels of frustration reported by the participants in the prediction+self-explanation group. These participants may have been less interested in the content due to their increased levels of frustration, in addition to their higher reported levels of mental

effort. A lack of significant differences related to interest may also be attributed to the nature of the questions from the instrument. The questions focused on interest in the content, specifically physics principles, rather than perceptions of the strategies employed. It appears, however, that the strategies in this study did not influence learners' interest in the content area of the instructional materials.

Limitations

Limitations to the present study should be recognized. An important threat to internal validity is the reliance on self-report data for measuring dimensions of cognitive load and interest. A direction for future research might be to obtain physiological measures or responses in think-aloud protocols. The short duration of treatment may have influenced results. While this study examined the effects of instructional strategies for learning principles during a single instructional session, future studies should increase the duration of the intervention to further examine effects.

External validity of this study may be threatened by the use of convenience sampling and limits the generalizability of the results to all populations of learners. However, the researcher sampled courses outside of science fields in order to minimize the chance that students would have prior knowledge regarding the content presented in the instructional materials. An additional threat to ecological validity is whether the results would generalize to all forms of instructional simulations. The simulation used in this study is one containing a simple graphical representation of user inputs and resulting simulation outputs. Simulations can vary greatly in how the underlying model is represented to the learner and the results of this study may not generalize to the use of more elaborate or complex simulations.

Implications

This research demonstrated that generative strategies can be a valuable instructional support for learning within a simulation environment. As compared with the control group, these participants exhibited improved learning and more accurate calibration judgments as compared with a recommended guided discovery approach. Paraphrasing proved to be an effective strategy that learners conducted with ease, reflected in several aspects of the outcome measures: (a) a greater breadth of idea units presented in paraphrases as compared with self-explanations, (b) moderate effort and demand levels, and (c) a high confidence level in their performance during the learning phase.

Prediction+self-explanation, while an effective strategy for learning and calibration (prediction) accuracy, led to higher levels of frustration and lower confidence levels in performance. There was, however, great variation in the quality of self-explanations within this group. This variability in quality influenced test performance and both prediction and postdiction judgments of test performance. Perhaps with either increased practice in this strategy or feedback on the quality or content of self-explanations an even greater improvement in learning and calibration accuracy could be obtained, without negatively affecting frustration and confidence in performance.

This research has attempted to provide an additional means of instructional supports for implementation with simulation learning to address many of the difficulties learners experience as noted by de Jong and van Joolingen (1998). Incorporating a guided discovery approach is a superior method of instruction over a pure discovery approach (Alfieri et al., 2011; Mayer, 2004). In addition to guided discovery, however,

instructional simulations should incorporate instructional strategies, such as paraphrasing or predicting and self-explaining, to encourage learners to engage with the content and generate meaning. These strategies should facilitate learning of principles in a variety of mediums (e.g. print, animations) including simulations.

Future research is needed to further explore the effects of generative strategies and learning, specifically with principles. In addition, future research should examine prediction+self-explanation over a longer period of time and the effects on mental effort, frustration, performance, and calibration accuracy. Additional research is needed regarding the effects of feedback on the quality of self-explanations, as well as the possible learner attributes that enabled novice learners to create deep self-explanations. Last, research should continue to examine the effects of generative strategies on the dimensions of cognitive load as this appears to be a promising line of research in instructional design.

Conclusions

Although different benefits were noted for the two generative strategies, this study provides a different approach to the body of research guided by cognitive load theory (Sweller et al., 2011). The majority of the existing research has focused on strategies to minimize the negative effects of extraneous cognitive load, or the manner in which instructional materials are presented, when dealing with instructional materials that have high intrinsic cognitive load (i.e. complexity), as well as presentation strategies for material with high intrinsic cognitive load. This study employed instructional strategies to direct attention to germane resources related to schema construction when learners were faced with materials of high intrinsic cognitive load. Additionally, this study

demonstrated the effectiveness of prediction+self-explanation for novice learners. The strategy of self-explanation has primarily been explored when learners with some level of prior knowledge studied worked examples (for a review, see Sweller et al., 2011). Consistent with findings from de Bruin et al., (2007), novice learners with no prior knowledge of the content area benefitted from self-explanation in this study.

Finally, this study has contributed to the research related to self-regulated learning. Participants benefited from the feedback they received through the generative strategies and were able to more accurately assess their understanding, an important characteristic of the self-regulated learner (Zimmerman, 2000). Little research exists examining the effects generative strategies have on calibration judgments, particularly with simulation learning. This study demonstrated that generative strategies are an effective means to improve performance prediction judgments and have the potential to affect postdiction judgments.

REFERENCES

Afflerbach, P. (1990). The influence of prior knowledge and text genre on readers' prediction strategies. *Journal of Literacy Research, 22*(2), 131-148.

Afflerbach, P., & Walker, B. (1990). Prediction instruction in basal readers. *Reading Research and Instruction, 29*(4), 26-45. doi: 10.1080/19388079009558022

Alessi, S., & Trollip, S. R. (2001). *Multimedia for learning: Methods and development.* (3rd ed.). Boston, MA: Allyn & Bacon.

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

Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. D. Pearson, R. Barr, M. L. Kamil & P. Modenthal (Eds.), *Handbook of reading research* (Vol. 1, pp. 255-291). White Plains, NY: Longman.

Antonenko, P., Paas, F., Grabner, R., & van Gog, T. (2010). Using Electroencephalography to Measure Cognitive Load. *Educational Psychology Review, 22*(4), 425-438. doi: 10.1007/s10648-010-9130-y

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215. doi: 10.1037/0033-295x.84.2.191

Bangert-Drowns, R. L., Kulik, J., & Kulik, C. (1985). Effectiveness of computer-based education in secondary schools. *Journal of Computer-Based Instruction, 12*, 59-68.

Bol, L., Hacker, D. J., O'Shea, P., & Allen, D. (2005). The influence of overt practice, achievement level, and explanatory style on calibration accuracy and performance. *The Journal of Experimental Education, 73*(4), 269-290.

Bol, L., Hacker, D. J., Walck, C. C., & Nunnery, J. A. (2012). The effects of individual or group guidelines on the calibration accuracy and achievement of high school biology students. *Contemporary Educational Psychology, 37*(4), 280-287. doi: 10.1016/j.cedpsych.2012.02.004

Bretzing, B. H., & Kulhavy, R. W. (1979). Notetaking and depth of processing. *Contemporary Educational Psychology, 4*(2), 145-153. doi: 10.1016/0361-476X(79)90069-9

Byrne, M. D., Catrambone, R., & Stasko, J. T. (1999). Evaluating animations as student aids in learning computer algorithms. *Computers & Education, 33*(4), 253-278. doi: 10.1016/S0360-1315(99)00023-8

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science, 1*, 73-105. doi: 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*(2), 145-182. doi: 10.1016/0364-0213(89)90002-5

Chi, M. T. H., De Leeuw, N., Chiu, M.-H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*(3), 439-477. doi: 10.1016/0364-0213(94)90016-7

Chi, M. T. H., & VanLehn, K. A. (1991). The content of physics self-explanations. *The Journal of the Learning Sciences, 1*(1), 69-105.

Collins, A., Brown, J. S., & Larkin, K. M. (1980). Inferences in text understanding. In R. J. Spiro, b. C. Bruce & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 385-407). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Craig, R. C. (1956). Directed versus independent discovery of established relations. *Journal of Educational Psychology, 47*(4), 223-234. doi: 10.1037/h0046768

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671-684.

Davey, B., & McBride, S. (1986). Effects of question-generation training on reading comprehension. *Journal of Educational Psychology, 78*(4), 256-262. doi: 10.1037/0022-0663.78.4.256

de Bruin, A. B. H., Rikers, R. M. J. P., & Schmidt, H. G. (2007). The effect of self-explanation and prediction on the development of principled understanding of chess in novices. *Contemporary Educational Psychology, 32*(2), 188-205. doi: 10.1016/j.cedpsych.2006.01.001

de Jong, T., & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68*(2), 179-201. doi: 10.3102/00346543068002179

Evans, J. L., & Homme, L. E. (1962). The Ruleg system for the construction of programmed verbal learning sequences. *The Journal of Educational Research, 55*(9), 513-518.

Freeman, R. H. (1982). *Improving the comprehension of stories using predictive strategies*. Paper presented at the Annual Meeting of the International Reading Association, Chicago, IL.

Gagne, R. M., & Brown, L. T. (1961). Some factors in the programming of conceptual learning. *Journal of Experimental Psychology, 62*(4), 313-321. doi: 10.1037/h0049210

Gerjets, P., Scheiter, K., & Catrambone, R. (2006). Can learning from molar and modular worked examples be enhanced by providing instructional explanations and prompting self-explanations? *Learning and Instruction, 16*(2), 104-121. doi: 10.1016/j.learninstruc.2006.02.007

Glover, J. A., Plake, B. S., Roberts, B., Zimmer, J. W., & Palmere, M. (1981). Distinctiveness of encoding: The effects of paraphrasing and drawing inferences on memory from prose. *Journal of Educational Psychology, 73*(5), 736-744. doi: 10.1037/0022-0663.73.5.736

Glover, J. A., Timme, V., Deyloff, D., Rogers, M., & Dinell, D. (1987). Oral directions: Remembering what to do when. *Journal of Educational Research, 81*(1), 33-40.

Hacker, D., Bol, L., & Bahbahani, K. (2008). Explaining calibration accuracy in classroom contexts: the effects of incentives, reflection, and explanatory style. *Metacognition and Learning, 3*(2), 101-121. doi: 10.1007/s11409-008-9021-5

Hacker, D. J., Bol, L., Horgan, D. D., & Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology, 92*(1), 160-170. doi: 10.1037/0022-0663.92.1.160

Hansen, J. (1981). The effects of inference training and practice on young children's reading comprehension. *Reading Research Quarterly, 16*(3), 391-417.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of experimental and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam:: North Holland.

Hausmann, R. G. M., & VanLehn, K. (2010). The effect of self-explaining on robust learning. *International Journal of Artificial Intelligence in Education, 20*(4), 303-332.

Hegarty, M., Kriz, S., & Cate, C. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition and Instruction, 2*(4), 325-360.

Jonassen, D. (1988). Integrating learning strategies into courseware to facilitate deeper processing. In D. Jonassen (Ed.), *Instructional designs for microcomputer courseware* (pp. 151-181). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors, 40*(1), 1-17. doi: 10.1518/001872098779480587

Kasmer, L., & Kim, O.-K. (2011). Using prediction to promote mathematical understanding and reasoning. *School Science and Mathematics, 111*(1), 20-33.

Keren, G. (1991). Calibration and probability judgements: Conceptual and methodological issues. *Acta Psychologica, 77*(3), 217-273. doi: 10.1016/0001-6918(91)90036-Y

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363-394.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75-86. doi: 10.1207/s15326985ep4102_1

Kittell, J. E. (1957). An experimental study of the effect of external direction during learning on transfer and retention of principles. *Journal of Educational Psychology, 48*(7), 391-405. doi: 10.1037/h0046792

Kulik, C.-L. C., & Kulik, J. A. (1991). Effectiveness of computer-based instruction: An updated analysis. *Computers in Human Behavior, 7*(1-2), 75-94. doi: 10.1016/0747-5632(91)90030-5

Kunkel, D., & Kovaric, P. (1983). *Mental effort and learning from TV: Comparing expectations for PBS and commercial network programming.* Paper presented at the International Communication Association, Dallas, TX.

Lewis, E. L., Stern, J. L., & Linn, M. C. (1993). The effect of computer simulations on introductory thermodynamics understanding. *Educational Technology, 33*(1), 45-58.

Maki, R. H., Foley, J. M., Kajer, W. K., Thompson, R. C., & Willert, M. G. (1990). Increased processing enhances calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(4), 609-616. doi: 10.1037/0278-7393.16.4.609

Markle, S. M. (1969). *Good frames and bad: A grammar of frame writing.* New York: John Wiley & Sons, Inc.

Mayer, R. E. (2004). Should there be a three strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist, 59*, 14-19.

Okada, T., & Simon, H. A. (1997). Collaborative discovery in a scientific domain. *Cognitive Science, 21*(2), 109-146. doi: 10.1207/s15516709cog2102_1

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and monitoring activities. *Cognition & Instruction, 1*, 117-175.

Peper, R. J., & Mayer, R. E. (1986). Generative effects of note-taking during science lectures. *Journal of Educational Psychology, 78*(1), 34-38. doi: 10.1037/0022-0663.78.1.34

Reigeluth, C. M., & Schwartz, E. (1989). An instructional theory for the design of computer-based simulations. *Journal of Computer-Based Instruction, 16*(1), 1-10.

Renkl, A. (1997). Learning from Worked-Out Examples: A Study on Individual Differences. *Cognitive Science, 21*(1), 1-29. doi: 10.1207/s15516709cog2101_1

Rigney, J. W., & Lutz, K. A. (1976). Effect of graphic analogies of concepts in chemistry on learning and attitudes. *Journal of Educational Psychology, 68*(3), 305-311. doi: 10.1037/0022-0663.68.3.305

Rivers, R. H., & Vockell, E. (1987). Computer simulations to stimulate scientific problem solving. *Journal of Research in Science Teaching, 24*(5), 403-415. doi: 10.1002/tea.3660240504

Rummelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Mahwah, NJ: Lawrence Erlbaum Associates.

Salomon, G. (1981). *Communication and education: Social and psychological interactions*. Beverly Hills, CA: Sage Publications.

Salomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology, 76*(4), 647-658. doi: 10.1037/0022-0663.76.4.647

Salomon, G., & Leigh, T. (1984). Predispositions about learning from print and television. *Journal of Communication, 34*, 119-135.

Schommer, M., & Surber, J. R. (1986). Comprehension-monitoring failure in skilled adult readers. *Journal of Educational Psychology, 78*(5), 353-357. doi: 10.1037/0022-0663.78.5.353

Schraw, G., Bruning, R., & Svoboda, C. (1995). Sources of situational interest. *Journal of Literacy Research, 27*(1). doi: 10.1080/10862969509547866

Sweller, J. (1999). *Instructional design in technical areas*. Camberwell, Victoria: The Australian Council for Educational Research Ltd.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review, 22*(2), 123-138. doi: 10.1007/s10648-010-9128-5

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer.

Thiede, K. W., & Anderson, M. C. M. (2003). Summarizing can improve metacomprehension accuracy. *Contemporary Educational Psychology, 28*(2), 129-160. doi: 10.1016/S0361-476X(02)00011-5

Tuovinen, J. E., & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of Educational Psychology, 91*(2), 334-341. doi: 10.1037/0022-0663.91.2.334

van Berkum, J. J. A., & de Jong, T. (1991). Instructional environments for simulations. *Education & Computing, 6*, 305-358.

White, R. T., & Gunstone, R. F. (1992). *Probing understanding*. London: Falmer.

Wittrock, M. C. (1974). Learning as a generative process. *Educational Psychologist, 19*(2), 87-95. doi: 10.1080/00461520903433554

Wittrock, M. C. (1979). The cognitive movement in instruction. *Educational Researcher, 8*(2), 5-11. doi: 10.3102/0013189X008002005

Wittrock, M. C. (1989). Generative processes of comprehension. *Educational Psychologist, 24*, 345-376. doi: 10.1207/s15326985ep2404_2

Wittrock, M. C. (1992). Generative learning processes of the brain. *Educational Psychologist, 27*(4), 531. doi: 10.1207/s15326985ep2704_8

Wittrock, M. C., & Alesandrini, K. (1990). Generation of summaries and analogies and analytic and holistic abilities. *American Educational Research Journal, 27*(3), 489-502.

Zimmerman, B. (1990). Self-Regulated learning and academic achievement: An overview. *Educational Psychologist, 25*(1), 3-17.

Zimmerman, B. J. (1986). Becoming a self-regulated learner: Which are the key subprocesses? *Contemporary Educational Psychology, 11*(4), 307-313. doi: 10.1016/0361-476X(86)90027-5

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts & P. R. Pintrich (Eds.), *Handbook of self-regulation* (pp. 13-39). New York: Academic Press.

Appendix A. Instructional Assignments

A projectile is an object upon which the only force acting is gravity. Examples of projectiles are objects dropped or thrown upwards. The following assignments will allow you to explore the path a projectile travels under various conditions.

First, you will discover how **launch angle** affects projectile distance and flight time.

Control	Paraphrase group	Predict+self-explain group
		How do you predict launch angle will affect the distance traveled? What about flight time?

Initial conditions				Document distance traveled	Document flight time
Ht	Angle	Veloc	Mass		
0	15°	10 m/s	10kg		
0	20°	10 m/s	10kg		
0	45°	10 m/s	10kg		
0	60°	10 m/s	10kg		
0	75°	10 m/s	10kg		

	In your own words, explain the relationship between angle and distance and flight time. Explain the results of the experiment by relating it to your own paraphrase.	How do the results of this experiment compare with your initial prediction? Explain why the results confirmed or disproved your initial prediction.

Next, we'll explore launch angles a little further.

Two angles are **complementary** when they add up to 90°. For example, 30° and 60° are complementary angles. Now let's see how complementary angles affect projectile distance.

Control	Paraphrase group	Predict+self-explain group
		How do you think the distance traveled will compare between two projectiles launched with complementary angles (e.g. 30° and 60°)?

Initial conditions:				Document distance traveled	Document flight time
Ht.	Angle	Veloc.	Mass		
0	25°	10 m/s	10kg		
0	65°	10 m/s	10kg		
0	30°	10 m/s	10kg		
0	60°	10 m/s	10kg		
0	42°	10 m/s	10kg		
0	48°	10 m/s	10kg		

	In your own words, explain the relationship between complementary angles and distance and flight time. Explain the results of the experiment by relating it to your own paraphrase.	How do the results of this experiment compare with your initial prediction? Explain why the results confirmed or disproved your initial prediction.

Now we'll examine how the **mass** of an object affects the travel path of a projectile.

Control	Paraphrase group	Predict+self-explain group
		How do you predict the mass of an object will affect the distance a projectile travels? What about the effect on how long the projectile is in the air? The maximum height the projectile achieves?

Initial conditions				Document distance traveled	Document flight time	Document height
Ht.	Angle	Veloc.	Mass			
0	30	10 m/s	10kg			
0	30	10 m/s	25kg			
0	30	10 m/s	75kg			
0	30	10 m/s	200kg			

Control	Paraphrase group	Predict+self-explain group
	In your own words, explain the relationship between mass and a projectile's distance, flight time, and height. Explain the results of the experiment by relating it to your own paraphrase.	How do the results of this experiment compare with your initial prediction? Explain why the results confirmed or disproved your initial prediction.

Initial velocity is an additional variable that affects the motion of a projectile.

Initial velocity affects the horizontal motion of a projectile.

Control	Paraphrase group	Predict+self-explain group
		How do you predict changing the initial velocity of a projectile will affect the distance an object travels? What about the flight time? And height?

Initial conditions				Document	Document flight	Document
Ht.	Angle	Veloc.	Mass	distance traveled	time	height
0	30	10 m/s	45kg			
0	30	15 m/s	45kg			
0	30	20 m/s	45kg			
0	30	25 m/s	45kg			

	In your own words, explain the relationship between initial velocity and a projectile's distance, flight time, and height. Explain the results of the experiment by relating it to your own paraphrase.	How do the results of this experiment compare with your initial prediction? Explain why the results confirmed or disproved your initial prediction.

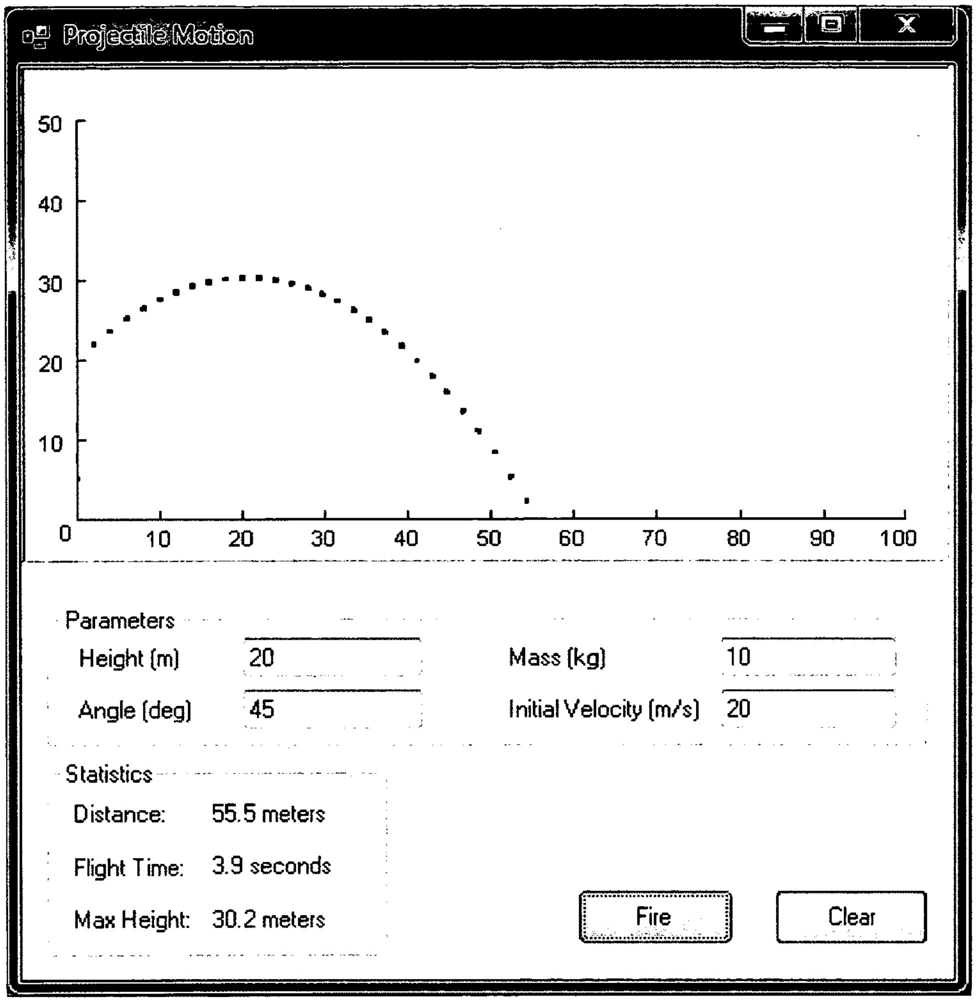
The effect of a projectile's **launch height** is the last variable we will explore.

Control	Paraphrase group	Predict+self-explain group
		How do you predict changing a projectile's launch height will affect the distance an object travels? What about the flight time? And height?

Initial conditions				Document distance traveled	Document flight time	Document height
Ht.	Angle	Veloc.	Mass			
0	45°	5m/s	10kg			
20m	45°	5m/s	10kg			
20m	43	5m/s	10kg			
20m	41	5m/s	10kg			
20m	39	5m/s	10kg			
20m	37	5m/s	10kg			

	In your own words, explain the relationship between launch height and a projectile's distance, flight time, and height. Explain the results of the experiment by relating it to your own paraphrase.	How do the results of this experiment compare with your initial prediction? Explain why the results confirmed or disproved your initial prediction.

Appendix B. Instructional Simulation



Appendix C. Achievement Test Blueprint

	Recall	Application	Evaluation	Near Transfer	Total
A projectile launched from the ground will obtain the maximum distance with a 45° launch angle	2	2	1	2	7
Same distance achieved with complementary angles.	2	2	1	2	7
Mass of a projectile does not affect distance or flight time.	2	2	1	1	6
As initial velocity increases, distance, height, flight time increase.	2	2	1	2	7
As launch height increases, optimum launch angle decreases for max distance.	2	2	1	2	7
	10	10	5	19	34

Appendix D. Achievement Test Items

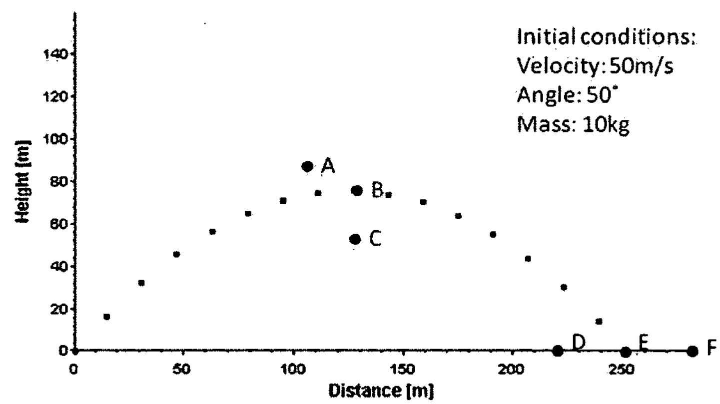
Your friend is playing with a slingshot that has a fixed launch angle and you notice that the balls appear to travel different distances with each release. What factor(s) might explain this discrepancy in distances?

Two projectiles are launched from a height of 15ft and with the same initial velocity. Projectile A is launched with an angle of 43° and Projectile B with an angle of 39°. Which of the following statements is correct?

a. Projectile A will have a longer flight time
b. Projectile A will travel a farther distance
c. The projectiles will have the same flight times
d. The projectiles will travel the same distance

If the initial velocity of a projectile is increased while other variables are kept the same, which of the following will occur?

a. The distance traveled increases but the height decreases
b. The distance traveled and the height increase
c. The distance traveled decreases but the height increases
d. The distance traveled and the height decrease



If the launch angle in the scenario was changed to 40°, the projectile would most likely travel through which point(s) on the graph?

a. A and D
b. B and E
c. B and F
d. C and E

Two projectiles are launched from a height of 15ft, one with an angle of 44° and another with an angle of 38°. How would the distances traveled by each projectile compare?

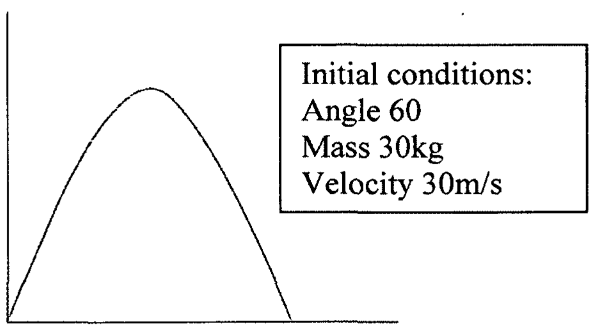
Two cars, one with a mass of 1500kg and another with a mass of 2000kg drive off a cliff with the same initial velocity and angle. Which will travel the farthest distance before landing?

a. 1500kg car
b. 2000kg car
c. There is not enough information given.
d. They will travel the same distance.

When a projectile is launched from the ground, what angle will lead to the farthest distance traveled?

A comparable distance may be achieved when launching projectiles with which of the following?
a. 30° and 15°
b. 35° and 60°
c. 40° and 50°
d. 45° and 15°

The home team has decided to kickoff at the start of a football game. The goal is to send the football as far as possible down the field. How should the coach advise the kicker in regards to angle and velocity?



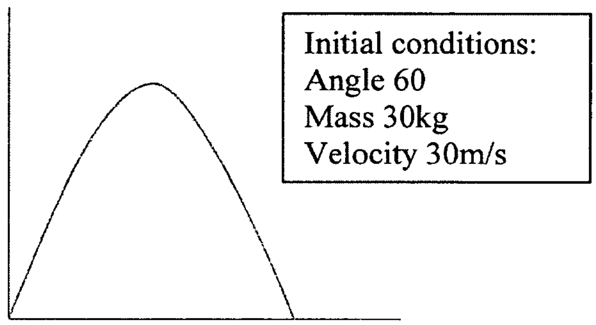
Which of the following would increase the distance of a projectile shown in the above image?
a. Decrease angle from 60° to 45°
b. Decrease velocity below 30 m/s
c. Increase angle from 60° toward 90°
d. Increase mass above 40kg

Two masses are launched from the ground. Projectile A is launched with an initial velocity of 20m/s and projectile B with an initial velocity of 25m/s. Which of the following statements is true?
a. Both projectiles will have the same flight times.
b. Projectile A will have a shorter flight time
c. Projectile A will reach a greater height
d. Projectile A will travel a farther distance

A shot is put (thrown) from above the athlete's shoulder level. The launch angle that will produce the longest range is less than 45°. Explain why.

A projectile with a mass of 30kg is launched from the ground at a 35° angle with an initial velocity of 15m/s achieves a distance of 85m. Which of the following would also result in a distance of 85m?

a. A launch angle of 10°, a mass of 55kg, and an initial velocity of 15 m/s
b. A launch angle of 10°, a mass of 30kg, and an initial velocity of 15m/s
c. A launch angle of 55°, a mass of 20kg, and an initial velocity of 15m/s
d. A launch angle of 55°, a mass of 25kg, and an initial velocity of 20m/s



Which of the following would increase the distance of a projectile shown in the above image?

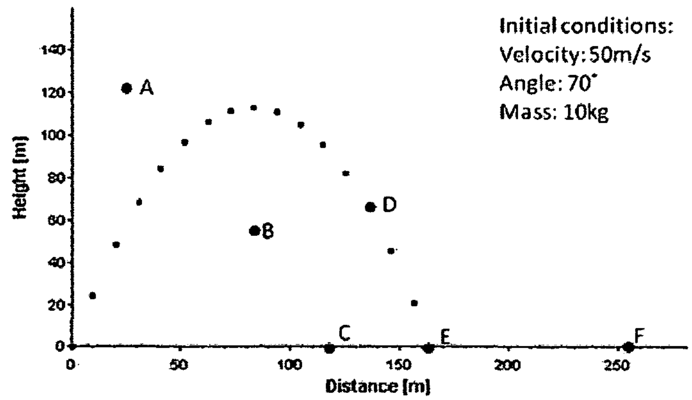
a. Decrease angle from 60° to less than 30°
b. Decrease mass below 20kg
c. Increase angle from 60° toward 90°
d. Increase launch height to 20ft

Two projectiles are launched from the ground, projectile A with a 30° angle and projectile B with a 45° angle. Which of the following statements is correct?

a. Both projectiles will travel the same distance
b. Projectile A will have a greater flight time
c. Projectile A will reach a greater height
d. Projectile A will travel a shorter distance

Which of the following does NOT affect the distance a projectile travels?

a. Initial velocity
b. Launch angle
c. Launch height
d. Projectile mass

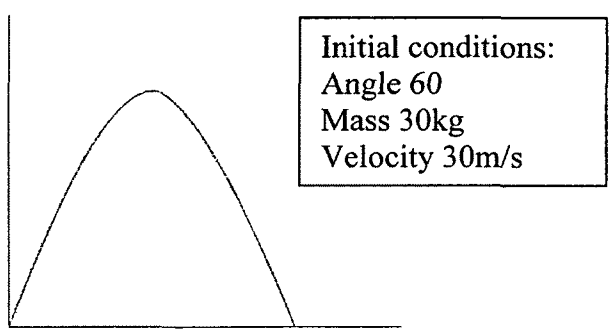


If the launch angle in this scenario is reduced to 45°, the projectile would most likely travel through which point(s) on the graph?

a. A and F
b. B and E
c. D and E
d. D and F

Launching projectiles with complementary angles results in which of the following?

a. Same projectile heights
b. Same projectile heights and distances
c. Same projectile distances
d. Same projectile time of flights



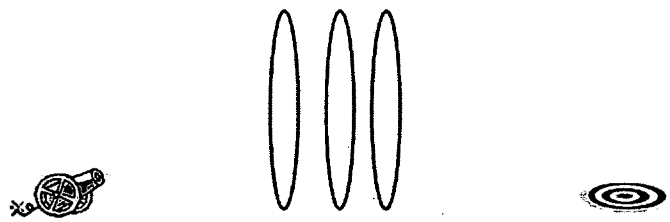
Which of the following would increase the distance of a projectile shown in the above image?

a. Decrease angle from 60° to less than 30°
b. Decrease velocity below 30 m/s
c. Increase mass above 40kg
d. Increase velocity above 30 m/s

Two masses are tossed with the same initial velocity. The heavier has twice the mass of the lighter. Which statement is correct?

a. The heavier mass flies twice as far as the lighter

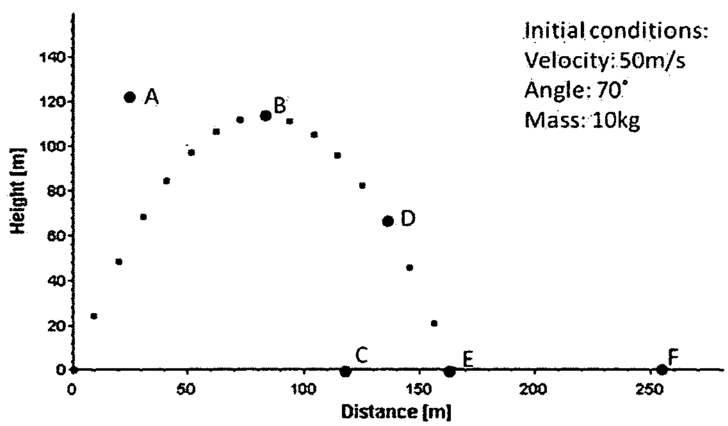
b. The heaver mass flies twice as far as the heavier
c. The lighter mass has a higher trajectory
d. The two masses have the same trajectory



A human cannonball is one of the attractions at the local circus. You have been charged with manning the cannon. The cannon has been placed the appropriate distance from its target and is currently angled at 22°. Your observations tell you that the performer won't clear the series of rings between the cannon and the target. What adjustments would you make?

When compared to other launch angles, a launch angle of 45° will result in which of the following:

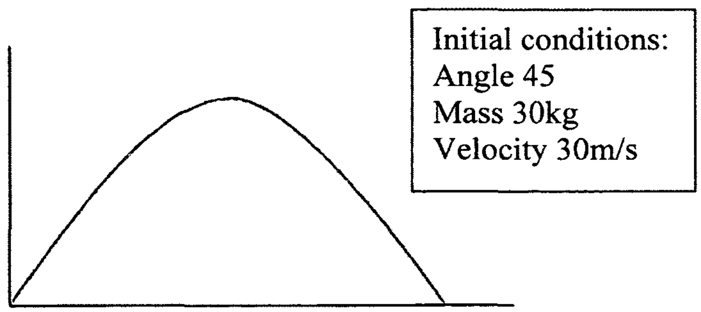
a. Farthest distance traveled
b. Highest projectile path
c. Lowest projectile path
d. Shortest distance traveled



If the mass of the projectile was increased to 30kg, the projectile would most likely travel through which point(s) on the graph?

a. A and C
b. A and E
c. B and E
d. D and F

When the launch height is increased what must be adjusted to obtain maximum distance?



Which of the following would decrease the flight time of a projectile thrown in the above image?

a. Decrease angle from 45° to less than 30°
b. Decrease mass below 20kg
c. Increase angle from 45° toward 90°
d. Increase mass above 40kg

Firefighters report on the scene for a fire occurring in a field. The hose isn't long enough to reach the field itself, so the firefighters attach the hose to a stand on the ground. How could the firefighters angle the stream of water to reach the field?

What effect does mass have on a projectile's distance traveled, flight time, and height?

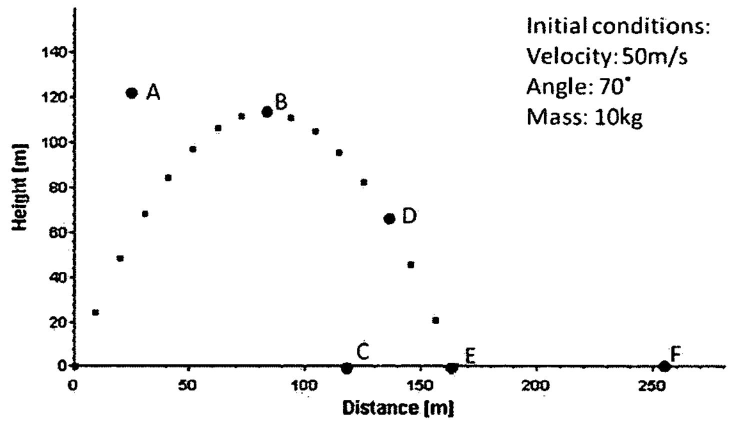
A pirate ship is moored 560m from a harbor. The canon will most certainly hit the pirate ship if angled at 63°. However, time is critical as the pirate ship is preparing to attack the harbor. How would you advise the commander of the fort?

A shorter flight time will occur when:

a. Initial velocity is reduced
b. Launch angle is increased
c. Launch height is increased
d. Mass is reduced

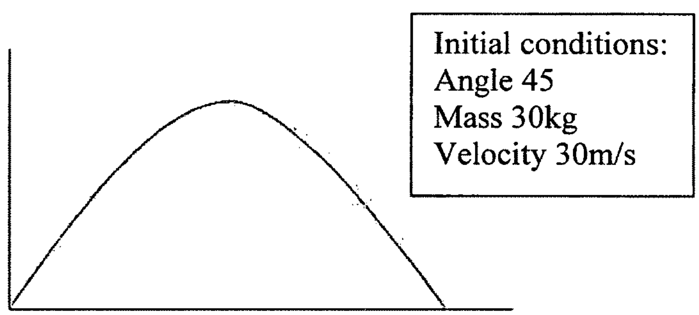
When the launch height is raised from ground-level to 10ft, how are the distance traveled and time of flight affected? Assume all other factors are kept constant.

In a long jump event, the jumper should strive for which angle?



If the initial velocity in the scenario is increased to 80m/s, the projectile would most likely travel through which point(s) on the graph?

a. A and F
b. A and C
c. B and E
d. D and F



Which of the following would decrease the flight time of a projectile thrown in the above image?

a. Decrease velocity below 30 m/s
b. Decrease mass below 20kg
c. Increase launch height to 20ft
d. Increase velocity above 30 m/s

You are responsible for training the new quarterback and kicker for a football team. How would you advise the quarterback on the angle he should throw as compared to the angle the kicker should kick?

a. Throw at a higher angle
b. Throw at a lower angle
c. Throw at the same angle
d. Try for a 45° angle

Appendix E. Generation Evaluation Rubric

	Present	Absent
Principle 1: A 45° angle results in the farthest distance traveled when a projectile is launched from the ground (primary). Angles greater than 45° result in an increased flight time and decreased distance and/or Angles lower than 45° result in a decreased flight time and decreased distance.		
Principle 2: Complementary angles result in the same distance traveled. Flight times differ when a projectile is launched with a complementary angle.		
Principle 3: Mass has no effect on a projectile's flight time, distance traveled, or projectile path height.		
Principle 4: An increase in velocity will result in an increased flight time, distance traveled, and projectile path height.		
Principle 5: When the launch height of a projectile is raised above ground level, an angle less than 45° results in the farthest distance traveled. An increase in the launch height of a projectile will result in an increased flight time, distance traveled, and projectile path height.		

Appendix F. Cognitive Load Questionnaire

Effort – repeated measure:

a. How hard did you have to work in your attempt to understand the contents of the learning environment?

0 (Very low)	100 (Very high)

Mental Demand:

a. How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)?

0 (Very low)	100 (Very high)

b. Was the learning task easy or demanding?

0 (Very low)	100 (Very high)

Performance:

a. How successful do you think you were in understanding the contents?

0 (Good)	100 (Very poor)

b. How satisfied were you with your performance in accomplishing the learning task?

0 (Good)	100 (Very poor)

Frustration Level:

a. How frustrated were you during the learning task?

0 (Very low)	100 (Very high)

Appendix G. Interest Questionnaire

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
I thought the instruction was very interesting					
I'd like to discuss this instruction with others at some point					
I would complete this instruction again if I had the chance					
I got caught-up in the instruction without trying to					
I'll probably think about the implications of this instruction for some time to come					
I thought the instruction's topic was fascinating					
I think others would find this instruction interesting					
I would like to learn more about this topic in the future					
The instruction was one of the most interesting things I've learned in a long time					
The instruction really grabbed my attention					

APPENDIX H. EXAMPLE ASSIGNMENTS

Control Group Example

In the following activities, you will examine the path of a projectile. Before you begin, review the instructions and example below.

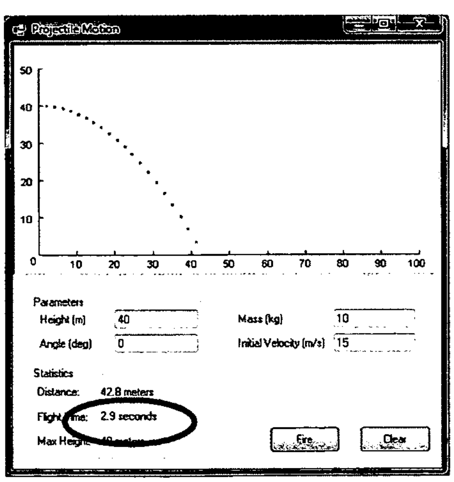
You will first be presented with a description of the task:

Let's examine the flight time for a dropped projectile as compared to a horizontally launched projectile.

You will use the simulation, launched after this instruction, to complete the experiment.

Set height	Set launch angle	Document flight time
20m.	-90° (dropped)	
20m.	0° (horizontally projected)	
30m	-90° (dropped)	
30m	0° (horizontally projected)	
40m	-90° (dropped)	
40m	0° (horizontally projected)	

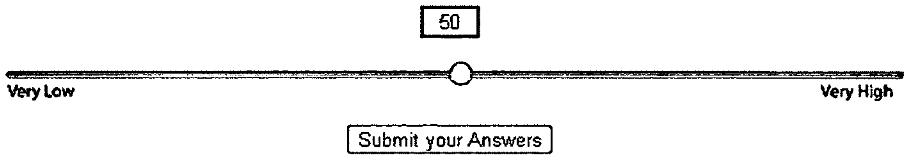
Enter the requested values from the simulation output of each launch in the table column on the right.



Set height	Set launch angle	Document flight time
20m.	-90° (dropped)	2.0s
20m.	0° (horizontally projected)	2.0s
30m	-90° (dropped)	2.5s
30m	0° (horizontally projected)	2.5s
40m	-90° (dropped)	2.9s
40m	0° (horizontally projected)	2.9s

After each activity, you will respond to a question by clicking your response on a sliding scale. Practice clicking your response in the example below. You can change your response before moving on.

How hard did you have to work in your attempt to understand the contents in the instruction?



If you are ready to begin with the instructional activity, click the next button below.

Paraphrase Example

In the following activities, you will examine the path of a projectile. Before you begin, review the instructions and example below.

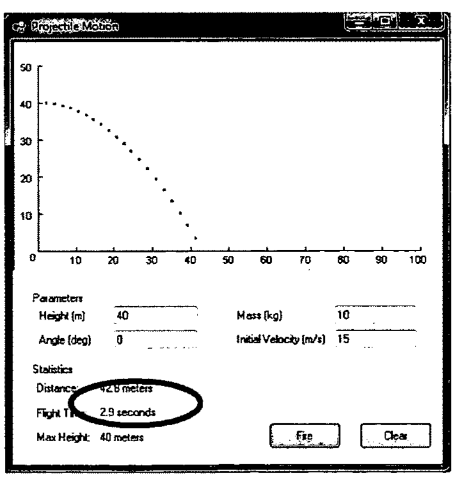
You will first be presented with a description of the task:

Let's examine the flight time for a dropped projectile as compared to a horizontally launched projectile.

You will use the simulation, launched after this instruction, to complete the experiment.

Set height	Set launch angle	Document flight time
20m.	-90° (dropped)	
20m.	0° (horizontally projected)	
30m	-90° (dropped)	
30m	0° (horizontally projected)	
40m	-90° (dropped)	
40m	0° (horizontally projected)	

Enter the requested values from the simulation output of each launch in the table column on the right.



Set height	Set launch angle	Document flight time
20m.	-90° (dropped)	2.0s
20m.	0° (horizontally projected)	2.0s
30m	-90° (dropped)	2.5s
30m	0° (horizontally projected)	2.5s
40m	-90° (dropped)	2.9s
40m	0° (horizontally projected)	2.9s

You will then be prompted to use the results from the experiment to explain the relationship found for the task.

In your own words, explain how flight times compare between dropped objects and horizontally projected objects.

It appears that the flight time is the same for a dropped object as compared with a horizontally projected object.

Notice the statement of the relationship

Last, you will be prompted to use the results from the experiment to explain the relationship found for the task.

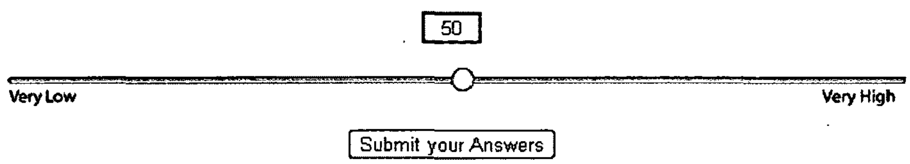
Explain the results of the experiment by relating it to your own paraphrase.

Objects that are dropped from a certain height have the same flight time as those that are horizontally projected. For example, when an object was dropped from a height of 20m, it had a flight time of 2.02s. The same flight time occurred when objects were projected horizontally.

Notice that supporting information from the experiment is provided to expand on the statement above.

After each activity, you will respond to a question by clicking your response on a sliding scale. Practice clicking your response in the example below. You can change your response before moving on.

How hard did you have to work in your attempt to understand the contents in the instruction?



If you are ready to begin with the instructional activity, click the next button below.

Prediction + self-explanation Example

In the following activities, you will examine the path of a projectile. Before you begin, review the instructions and example on the next page.

You will first be presented with a description of the task:

Let's examine the flight time for a dropped projectile as compared to a horizontally launched projectile.

You will be prompted to enter a prediction related to the task:

How do you predict the flight times will compare between a dropped and horizontally projected object from the same height?

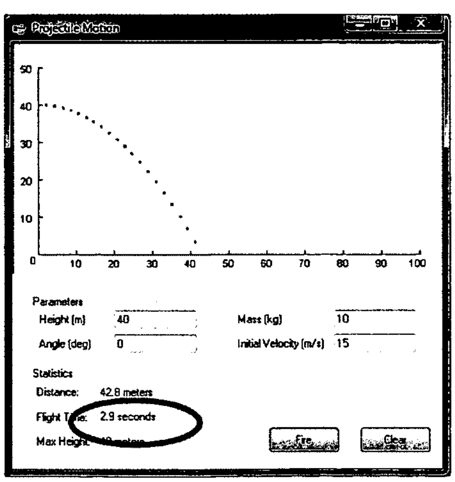
I think a dropped object will have a shorter flight time than one projected horizontally.

There is no right or wrong answer here. Just put what you think will happen.

You will use the simulation, launched after this instruction, to complete the experiment.

Set height	Set launch angle	Document flight time
20m.	-90° (dropped)	
20m.	0° (horizontally projected)	
30m	-90° (dropped)	
30m	0° (horizontally projected)	
40m	-90° (dropped)	
40m	0° (horizontally projected)	

Enter the requested values from the simulation output of each launch in the table column on the right.



Set height	Set launch angle	Document flight time
20m.	-90° (dropped)	2.0s
20m.	0° (horizontally projected)	2.0s
30m	-90° (dropped)	2.5s
30m	0° (horizontally projected)	2.5s
40m	-90° (dropped)	2.9s
40m	0° (horizontally projected)	2.9s

You will then be prompted to use the results from the experiment to compare with your initial prediction.

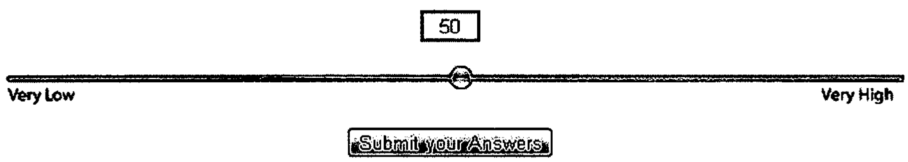
How do the results of this experiment compare with your initial prediction? Explain why the results confirmed or disproved your initial prediction.

I thought that a horizontally projected object would be in the air longer because the distance traveled would increase the flight time. In the experiment, I found that the dropped object and horizontally projected object have the same flight time. I think this is because gravity pulls both downward at the same rate.

Notice:
1. a comparison with the prediction.
2. a description of the experiment results and
3. a possible explanation why.

After each activity, you will respond to a question by clicking your response on a sliding scale. Practice clicking your response in the example below. You can change your response before moving on.

How hard did you have to work in your attempt to understand the contents in the instruction?



If you are ready to begin with the instructional activity, click the next button below.

VITA

Jennifer R. Morrison

EDUCATION

Old Dominion University, PhD candidate
Instructional Design and Technology (dissertation defended May 2013)
Dissertation: The effects of generative strategies on learning, cognitive load, and calibration accuracy in a simulation environment

Old Dominion University, M.S.
Occupational/Technical Studies, Instructional Design emphasis (2010)

Central Michigan University, B.M.
Music performance (2001)
Graduated magna cum laude

PUBLICATIONS

Morrison, G. R. and Morrison, J. R. (2013) Dual coding theory. In R. C. Richey (Ed.). *Encyclopedia of terminology of educational communications and technology* (pp. 99). New York: Springer.

Morrison, J. R., Watson, G. S., and Morrison, G. R. (2012) Comparison of restricted and traditional discussion boards on student critical thinking. *Quarterly Review of Distance Education, 13*(3) 167-176.

Morrison, J. R. and Greenwell, S. (2012). Classic articles in instructional design and distance education. In L. Moller and J. Huett (Eds.) *The next generation of distance education: Unconstrained learning* (pp. 251-258). New York: Springer.

Morrison, J. R. (2012). A review of *Instructional design for teachers: Improving classroom practice* by A. A. Carr-Chellman. *Tech Trends, 56,* 44-45.

PRESENTATIONS

Morrison, J. R., Watson, G. S., and Morrison, G. R. (2013, April) *Learning from diagrams and text in authentic science instruction.* Paper presentation at the Old Dominion University Graduate Student Achievement Day, Norfolk, VA.

Sandberg, H. M., Ringleb, S., Watson, G., Morrison, J. Schwartz, K., Deutsch, M., & Raymer, A. M. (2012, November). *Rehabilitation of listening in noise: A case study in aphasia.* Poster presented at the annual convention of the American Speech-Language-Hearing Association, Atlanta, GA.

Morrison, G. R., Watson, G. S., Robison, D. G., Morrison, J. R., Maddrell, J., Desmarais, R. M., Shirey, F. E., & Peck, L. (2012, October). *Equivalency theory in distance education in the age of globalization.* Paper presented at the annual meeting of the Association for the Educational Communications and Technology, Louisville, KY.

Morrison, G. R. and Anglin, G. J., & Morrison, J. R. (2012, October). *Redundancy with text and pictures: A contradiction.* Paper presented at the annual meeting of the Association for the Educational Communications and Technology, Louisville, KY.

Morrison, J. R., Robison, D. G., Martin, M. W., & Watson, G. S. (2012, April). *Cognitive load in a simulation for mooring a Coast Guard cutter.* Paper presented at the Modeling, Simulation, & Visualization Student Capstone Conference, Suffolk, VA.

Morrison, J. R., Watson, G. S., & Morrison, G. R. (2011, November). Comparison of moderated and non-moderated discussion board on student critical thinking. Paper presented at the annual meeting of the Association for Educational Communications and Technology, Jacksonville, FL.

Morrison, J. R., Robison, D. G., Martin, M.W., & Watson, G.S. (2011, November). *Managing cognitive load in training a complex multivariate task.* Paper presented at the annual meeting of the Association for Educational Communications and Technology, Jacksonville, FL.

Robison, D. G., Morrison, J. R., Martin, M.W., & Watson, G. S. (2011, November). *Managing cognitive load with a two-dimension ship mooring simulator.* Poster presented at the annual meeting of the Association for Educational Communications and Technology, Jacksonville, FL.

Sandberg, H., Ringleb, S., Watson, G., Deutsch, M., Morrison, J.R., & Raymer, A. (November, 2011) Impact of visual information on listening in noise in aphasia. Poster presented at the annual convention of the American Speech-Language-Hearing Association, San Diego, CA.

Morrison, J. R., 7 Watson, G. S. (2010, April). *Designing physical instructional simulations to promote learning from multiple external representations.* Paper presented at Virginia Modeling, Analysis, and Simulation Center Student Capstone Conference, Norfolk, VA.

RESEARCH IN PROGRESS

Morrison, J. R., Watson, G. S., and Morrison, G. R. *Exploring the redundancy effect in print-based instruction.* Manuscript submitted for journal review.

Morrison, G. R., Anglin, G. J., and Morrison, J. R. *Redundancy with text and pictures: A contradiction.* Manuscript in preparation.

Morrison, J. R., Watson, G. S., and Morrison, G. R. (November, 2013) *An examination of the redundancy effect in print-based instruction.* Featured research paper accepted for presentation at the annual meeting of the Association for Educational Communications and Technology, Anaheim, CA.

Morrison, J. R., Watson, G. S., and Bol, L. (November, 2013).*Simulation design strategies to promote knowledge construction and self-regulated learning.* Paper accepted for presentation at the annual meeting of the Association for Educational Communications and Technology, Anaheim, CA.

PROFESSIONAL EXPERIENCE

Teaching Assistance 2013
Old Dominion University, Norfolk, VA

Assistance with course design for an asynchronous graduate distance education course in Human Performance Assessment
Development and implementation of instructional units
Facilitation of weekly synchronous discussions

Graduate Research Assistant 2009 – Present
Virginia Modeling, Analysis, and Simulation Center of Old Dominion University, Suffolk, VA

Videotaped, edited, and verified multimedia stimuli to be administered in an experimental study
Recruited, scheduled, consented, and collected data with human subjects
Created game storyboards and scripts for traumatic brain injury rehabilitation project

Store Manager, Regional Credit Card Captain 2005 – 2009
J.Crew, Ann Arbor, MI

Responsible for mentoring all stores in region to properly train managers and associates, increasing performance in new credit card accounts
Mentored stores in market on college outreach initiative
Created training materials and credit card initiative; used company-wide
Assisted in development of new manager training

Co-Sales Manager 2004 – 2005
New York & Company, Lansing, MI

Created credit tracking program that was implemented throughout the district

Created register training program for district

Store Manager 2001 –2004

Gap, Inc., Lansing, MI

Mentored area management teams on markdown execution and merchandising during 2003

Responsible for the successful implementation of all training

PROFESSIONAL CONSULTATION AND INSTRUCTIONAL MATERIALS

Class-size study. Army Educational Advice Committee, TRADOC, 2013.

Technology Integration Training for Teachers, Richmond Preparatory Christian Academy, Richmond, VA. January, 2013.

Namibian upwelling systems. Instructional unit. Namibia, Africa: Namibian Dolphin Project, 2011. Instructional designer.

AWARDS AND HONORS

2013	Inducted into the Honor Society of Phi Kappa Phi
2011-2013	Graduate Research PhD Assistantship - Cognitive Learning Lab at the Virginia Modeling, Analysis, and Simulation Center (VMASC) of Old Dominion University.
2012	Best Presentation in Training & Education Track Award for *Cognitive load in a simulation for mooring a Coast Guard cutter.* Presented at the Modeling, Simulation, & Visualization Student Capstone Conference, Suffolk, VA, April 2012. With Robison, D.G., Martin, M.W., and Watson, G.S.
2009-2010	Graduate Research Master's Assistantship - Cognitive Learning Lab at the Virginia Modeling, Analysis, and Simulation Center (VMASC) of Old Dominion University.
2009	Association for Educational Communications and Technology Master's Scholarship Award.
1999-2001	School of Music Scholarship, Central Michigan University.
1998	Inducted into the Phi Et Sigma National Honor Society
1998	Inducted into the Golden Key International Honor Society

SERVICE AND LEADERSHIP

Old Dominion University

2013	Instructor for Student Virginia Education Association STEM Day
2012-2013	Organized and led research group for current Instructional Design and Technology PhD students to discuss the status of research initiatives, and to provide accountability and support through various stages of research projects.
2011, 2012	Assisted in development and implementation of Instructional Design and Technology new student orientation.

Association for Educational Communications and Technology (AECT)

2012	Proposal reviewer and session facilitator
2011	Proposal reviewer and session facilitator
2009	Volunteer and session facilitator

PROFESSIONAL ORGANIZATIONS

Association for Educational Communications and Technology (AECT)
American Educational Research Association (AERA)