Winter 2013

In Search for Instructional Techniques to Maximize the Use of Germane Cognitive Resources: A Case of Teaching Complex Tasks in Physics

Yekaterina Sliva
Old Dominion University

Recommended Citation

Sliva, Yekaterina. "In Search for Instructional Techniques to Maximize the Use of Germane Cognitive Resources: A Case of Teaching Complex Tasks in Physics" (2013). Doctor of Philosophy (PhD), dissertation, STEM and Professional Studies, Old Dominion University, DOI: 10.25777/ff5p-4c84
https://digitalcommons.odu.edu/stemps_etds/100

IN SEARCH FOR INSTRUCTIONAL TECHNIQUES TO MAXIMIZE THE USE OF GERMANE COGNITIVE RESOURCES: A CASE OF TEACHING COMPLEX TASKS IN PHYSICS

by

Yekaterina Sliva
Combined B.S. and M.S. January 1990, Moscow State 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

December 2013

Approved by:

Gary R. Morrison (Director)

Ginger S. Watson (Member)

Alexander L. Godunov (Member)

ABSTRACT

IN SEARCH FOR INSTRUCTIONAL TECHNIQUES TO MAXIMIZE THE USE OF GERMANE COGNITIVE RESOURCES: A CASE OF TEACHING COMPLEX TASKS IN PHYSICS

Yekaterina Sliva
Old Dominion University, 2013
Director: Dr. Gary R. Morrison

The purpose of this study was to introduce an instructional technique for teaching complex tasks in physics, test its effectiveness and efficiency, and understand cognitive processes taking place in learners' minds while they are exposed to this technique. The study was based primarily on cognitive load theory (CLT). CLT determines the amount of total cognitive load imposed on a learner by a learning task as combined intrinsic (invested in comprehending task complexity) and extraneous (wasteful) cognitive load. Working memory resources associated with intrinsic cognitive load are defined as germane resources caused by element interactivity that lead to learning, in contrast to extraneous working memory resources that are devoted to dealing with extraneous cognitive load. However, the amount of learner's working memory resources actually devoted to a task depends on how well the learner is engaged in the learning environment. Since total cognitive load has to stay within limits of working memory capacity, both extraneous and intrinsic cognitive load need to be reduced. In order for effective learning to occur, the use of germane cognitive resources should be maximized. In this study, the use of germane resources was maximized for two experimental groups by providing a learning environment that combined problem-solving procedure with prompts to self-explain with and without completion problems.

The study tested three hypotheses and answered two research questions. The first hypothesis predicting that experimental treatments would reduce total cognitive load was not supported. The second hypothesis predicting that experimental treatments would increase performance was supported for the self-explanation group only. The third hypothesis that tested efficiency measure as adopted from Paas and van Merriënboer (1993) was not supported. As for the research question of whether the quality of self-explanations would change with time for the two experimental conditions, it was determined that time had a positive effect on such quality. The research question that investigated learners' attitudes towards the instructions revealed that experimental groups understood the main idea behind the suggested technique and positively reacted to it. The results of the study support the conclusions that (a) prompting learners to self-explain while independently solving problems can increase performance, especially on far transfer questions; (b) better performance is achieved in combination with increased mental effort; (c) self-explanations do not increase time on task; and (d) quality of self-explanations can be improved with time. Results based on the analyses of learners' attitudes further support that learners in the experimental groups understood the main idea behind the suggested techniques and positively reacted to them. The study also raised concern about application of efficiency formula for instructional conditions that increase both performance and mental effort in CLT. As a result, an alternative model was suggested to explain the relationship between performance and mental effort based on Yerkes-Dodson law (1908).

Keywords: instructional design, cognitive load, complex tasks, problem-solving, self-explanation

COPYRIGHT PAGE

This dissertation is dedicated to my family who was very patient with me and supported my efforts throughout this journey.

ACKNOWLEDGMENTS

I have been lucky to be surrounded by a great group of scholars, researchers, and professionals who have helped me throughout the whole process of shaping this study and implementing it in a real classroom. Dr. Morrison was there for me providing mentoring and always challenging me to think deeper and prepare every small detail to make this study meaningful. Dr. Watson provided ongoing counseling on methodology. She was the first person I have shared my findings with, and she was the one who guided me to bring them together and come up with interpretations. Dr. Godunov's support helped me to believe in what I was doing. I am especially grateful for his collaboration on preparing materials for this study. Dr. Godunov was also the one who provided the environment where I could have access to a wonderful group of students. Finally, I want to acknowledge and thank students who participated in this study. They took time to do all the required work, especially those who approached me to get clarity on how to do it better and continued to improve their problem-solving skills throughout the duration of this study. Without their hard work, this study would not be possible.

TABLE OF CONTENTS

Page

Chapter

APPENDICES

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

Introduction

Cognitive skill acquisition is a process targeting the ability to solve problems in intellectual tasks, which can differ by problem complexity. Complex learning requires more than mastering isolated skills and foremost deals with learning to coordinate separate skills that are necessary for authentic real-life task performance. Moreover, integration of knowledge structures, problem-solving techniques, reasoning skills, and attitudes is critical for effective performance in complex learning (van Merriënboer, Kester, & Paas, 2006).

Cognitive Processes and Complex Tasks

Effective instructional methods for practicing complex tasks differ from those effective for simple tasks (van Merriënboer et al., 2006). In contrast to simple tasks, complex tasks typically impose a very high cognitive load on the learner's cognitive system. Building expertise in complex conceptual domains of knowledge, such as in geometry, physics, biology, or computer programming, can be explained through understanding of cognitive processes taking place in learners' working and long-term memory. Working memory capacity is considered limited to seven plus or minus two elements or chunks of information (Miller, 1956) with information lost after approximately 20 seconds (Peterson & Peterson, 1959), if not stored in a long-term memory or constantly refreshed. Information has to pass through working memory that has partly independent processing channels for visual/spatial and auditory/verbal information before it can be transferred to long-term memory that has virtually unlimited capacity. Since working memory is the gateway to long-term memory, the limitation of working memory should be the first consideration for the design of instruction. Learning is considered to take place through schema construction and automation, therefore the second consideration for the design of instruction is the construction and automation of schemas. Schemas bring together multiple information elements that can be treated as a single element when recalled in working memory. Schemas are constructed by either creating a new schema or by extending and modifying existing schemas. Schemas become automated with practice allowing both fluid performance on familiar tasks, and by freeing working memory capacity, performance on unfamiliar tasks that otherwise would be impossible.

Self-Regulated Learning

According to Zimmermann (1989), self-regulated learning refers to how students become masters of their own learning processes. Zimmermann explained that self-regulated learning is not a mental ability or a performance skill, but rather it is the self-directed process through which abilities are transformed into task-related skills; it involves active learning in terms of metacognition, motivation, and action control. Learners engaged in self-regulated learning have a clear understanding of how and why to employ self-regulatory strategies in order to acquire knowledge.

The facilitation of self-regulated learning is a balancing act between necessary external support and desired internal regulation (Koedinger & Aleven, 2007). From an instructional point of view, there are two ways to externally support self-regulated learning within problem-solving processes. Direct external support (i.e., direct instruction) facilitates explicit problem-solving strategies, their application, and transfer; and an indirect external support facilitates application of already existing problem-solving skills. For example, if learners already possess certain problem-solving strategies but fail to use this knowledge in a specific situation, it would be reasonable to prompt them to apply their existing strategic knowledge effectively. An instructional method for indirect support of the regulation of learners' problem-solving processes is prompting (Wirth, 2009). The purpose of prompts is to direct learners to perform a specific activity which is contextualized within a particular problem-solving situation (Davis, 2003). According to Davis (2003), prompts can be categorized as generic or directed. While generic prompts ask learners to reflect on their performed problem-solving activities, directed prompts provide them with an expert model of thinking in the problem-solving process. Therefore, from self-regulated learning perspective, the goal of an instructional strategy intervention is to introduce learners to specific instructional strategies that assist in task completion and support learners' self-regulated engagement in tasks so that they can learn to manage their cognitive processes during learning. The focus of this study was on building expertise in physics. In particular, this study investigated management of cognitive load during the process of skill acquisition in complex tasks in physics by means of prompts used to guide learners in their problem-solving activities.

Literature Review

Recent instructional design approaches tend to focus on authentic tasks that are based on complex real-life experiences. Such tasks require learners to employ and integrate their knowledge, skills, and attitudes necessary for effective task performance. However, complex tasks pose high cognitive load on learner's cognitive system that, in turn, may interfere with learning. Therefore, it is important to integrate knowledge about human cognitive architecture into the design of instruction. Cognitive load theory

(Sweller, 1988) provides designers with an important perspective for choosing appropriate instructional methods based on implications of task complexity relative to the learner's cognitive system.

Cognitive Load Theory

Cognitive load theory (Sweller, 1988) uses interactions between information structures and knowledge about cognition to design instruction. The theory emphasizes that working memory capacity is limited when dealing with novel information obtained through sensory memory. Cognitive load refers to the processing demands placed on working memory at a specific point in time.

Recent changes in cognitive load theory. Cognitive load theory (CLT) traditionally differentiated between three types of cognitive load: intrinsic, extraneous, and germane; and assumed that these three types of cognitive load are additive. Working-memory load may be affected by the complexity of the learning task (intrinsic cognitive load) or by the manner in which it is presented (extraneous cognitive load). Historically, germane cognitive load was viewed as the remaining working memory capacity, which was used for schema construction (Sweller, van Merriënboer, & Paas, 1998); this type of load was thought to occur when learners engage in a deep information processing such as mentally organizing the material and relating it to prior knowledge (DeLeeuw & Mayer, 2008).

Recently, Sweller, Ayers, and Kalyuga (2011) proposed to differentiate between two types of cognitive load imposed by instructional materials: intrinsic (useful) and extraneous (wasteful). On the other hand, the authors suggested that working memory resources can be divided into two types of resources: germane (i.e., resources devoted to intrinsic cognitive process) and extraneous (i.e., resources that deal with extraneous cognitive load). Sweller et al. (2011) explained that learner's working memory could be overloaded if the combined intrinsic and extraneous cognitive load exceeds its capacity. Sweller and his colleagues emphasized that germane cognitive load is a reflection of the amount of load imposed by intrinsic element interactivity and does not independently contribute to total cognitive load, thus, they have started using the term germane resources rather than germane load. Working memory resources that are actually devoted to dealing with intrinsic cognitive load and lead to meaningful learning are defined as germane resources.

According to Sweller et al. (2011), the primary goal of CLT is to devise instructional procedures that reduce extraneous cognitive load and thus decrease the working memory resources that need to be devoted to processing information that is extraneous to learning. The working memory resources that are not needed to deal with extraneous cognitive load can be redirected to deal with intrinsic cognitive load that is germane to the learning process. More efficient and effective learning can be achieved by eliminating or minimizing cognitive activities that are not essential for learning because they generate unnecessary load (i.e., extraneous cognitive load) typically caused by inappropriate instructional formats; and by managing essential for learning load (i.e., intrinsic cognitive load) determined by interacting elements of information. Intrinsic load should either be reduced or increased depending on available cognitive resources and instructional goals.

Some of the techniques recommended for managing intrinsic load on the initial stages of learning are segmenting learning tasks into smaller parts causing the learner to process less information at a time; pre-training learners in essential definitions and procedures prior to the main instructional session; and learning a limited number of selected isolated elements of information during an initial stage of instruction followed by the next stage of instruction that includes all interactive elements of information in their full complexity (Sweller et al., 2011).

Two approaches to managing intrinsic cognitive load on later stages of skill acquisition include self-explanations and varying the content or examples in the instruction. Intrinsic cognitive load could be productively increased by prompting students to self-explain problem-solving steps and procedures using their knowledge of domain principles (self-explanation effect; e.g., Renkl & Atkinson, 2003). A second strategy is to vary the content of learning task by considering different situations and conditions rather than similar ones (variability of worked examples; e.g., Paas & van Merriënboer, 1994).

Types of Cognitive Load: A Closer Look. *Intrinsic cognitive load* traditionally refers to the amount of cognitive processing required to comprehend material and depends on the number of elements of information that must be processed simultaneously and their interactivity (Clark, Nguyen, & Sweller, 2006). Sweller and Chandler (1994) explained that the complexity of the instruction increases when instructional content is composed of component parts or "elements" and there is a relationship between these elements (i.e., the elements "interact" with each other). Sweller and Chandler described this phenomenon as element interactivity. Similarly, when van Merriënboer and Sweller (2005) described element interactivity they noted that if the number of elements that need to be organized in the working memory increases linearly, then the number of their possible combinations increases exponentially. Thus, problems or content with high element interactivity are more difficult to understand as they may overwhelm working memory with intrinsic load and prevent the formation of a schema. Element interactivity can be determined by the number of interacting elements that the learner has to process at a particular level of expertise (van Merriënboer & Ayers, 2005). For example, learning the alphabet has low intrinsic load as learning A is not dependent on learning G. In contrast, solving a math story problem typically has high element interactivity as the learner must keep several interacting elements in working memory to solve the problem. Schemas that are stored in a long-term memory allow learners to process multiple elements as one element and decrease working memory load. Since intrinsic cognitive load depends on the complexity of the content, it was originally thought impossible to alter by instructional intervention (Sweller & Chandler, 1994). However, recent research suggests that this type of load can be reduced (Pollock et al., 2002; van Merriënboer et al., 2003).

Extraneous cognitive load is controllable and depends on the instructional intervention; in particular, extraneous load is determined by the design of the instruction. The reduction in extraneous cognitive load is critical when instructions contain materials that pose high intrinsic load (Sweller & Chandler, 1994; Paas, Renkl, & Sweler, 2003; van Merriënboer & Sweller, 2005). A combination of high intrinsic and high extraneous cognitive load may be detrimental to learning because working memory may be overloaded. If, in contrast, the intrinsic cognitive load is low due to low element interactivity, a high extraneous cognitive load due to poor design features may be less harmful. Total cognitive load has to stay within working memory limits. The reduction in extraneous cognitive load becomes the initial focus of the design of instruction when intrinsic cognitive load is high, as it is the easiest to control. This reduction in total cognitive load allows for additional working memory resources that are germane to learning to be used for schema development.

Reducing Extraneous Cognitive Load. The following techniques for the reduction in extraneous load during information presentation have been extensively studied: using integrated text and diagram formats instead of split-source formats (split-attention effect; e.g., Chandler & Sweller, 1991), avoiding presentation of redundant information (redundancy effect; e.g., Chandler & Sweller, 1991), and the use of multiple modalities to present mutually referring textual and pictorial information (modality effect; e.g., Mousavi, Low, & Sweller, 1995). In addition, to support acquisition of problem-solving skills, extraneous cognitive load can be reduced by presenting worked examples in integrated format prior to practicing problem-solving (worked examples effect; e.g., Sweller, 1999).

Another instructional technique for reducing extraneous cognitive load is the use of *completion problems*. Van Merriënboer and Krammer (1987) first suggested the use of completion problems to increase the transfer of computer programming skills. Completion problems are problems for which a given state, a goal state, and a partial solution are provided to learners who must complete that partial solution by providing intermediate steps. Completion problems are known to bridge worked examples and conventional problems. The completion problem effect indicates that solving completion problems yields higher transfer of acquired skills than conventional problem solving. An explanation for this effect is that learners who work on conventional problems apply means-ends analysis that poses high extraneous cognitive load on learners' working (Sweller, 1988). In contrast, while learners work on completion problems they focus their attention on problem states and associated solution steps enabling them to induce cognitive schemas, in particular cognitive schemas that allow for transfer of acquired skills (van Merriënboer, Shuurman, de Crook, & Paas, 2002). However, most of the studies on completion problems provided strong support to the completion problems effect but did not collect data on cognitive load (van Merriënboer, 1990; van Merriënboer & de Croock, 1992). Paas (1992) first compared the effects of completion problems, worked examples, and conventional problems on cognitive load during transfer test performance and training performance. Paas found that completion problems or worked examples required the same amount of mental effort during training and led to higher transfer test performance, combined with lower cognitive load during the test than conventional problems.

Maximizing the Use of Germane Cognitive Resources

A more recent development in the design of instructions based on CLT considerations is the employment of practices that maximize the use of germane cognitive resources (Paas, et al., 2003; Sweller et al., 1998; van Merriënboer & Sweller, 2005; Clark, et al., 2006). When extraneous and intrinsic cognitive loads are lowered, learners may have cognitive capacity freed that can be invested in processes that directly contribute to learning (i.e., germane cognitive resources). However, learners are unlikely to engage in such activities spontaneously (Renkl, Stark, Gruber, & Mandl, 1998; Renkl, 1999), therefore research efforts should be directed toward identifying instructional techniques that stimulate learners to invest cognitive resources in activities relevant for learning (van Gog & Paas, 2008). One of the methods to induce or activate germane resources is to engage learners in self-explanation activity (Clark et. al., 2006). Prior studies have established the advantages of self-explanation activity with respect to learning outcomes, however, these studies were mostly concerned with self-explanation activity during the study of worked examples (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Renkl, 1997; Renkl, et al., 1998; Renkl & Atkinson, 2003; Atkinson, Renkl, & Merrill, 2003). Aleven and Koedinger (2002) suggested that prompting to self-explain during problem solving rather than during example study also fosters learning. Aleven and Koedinger reported that problem solving in intelligent tutoring environment can be enhanced by prompting learners to self-explain by identifying the underlying problem-solving principles.

Self-explanation. Chi et al. (1989) found that an instructional strategy requiring students to generate and articulate explanations of their own reasoning or understanding enhances deeper learning. Self-explanations involve generating comments that contain domain-relevant information and provide links beyond the information given. For example, the instruction might prompt learner to use self-explanatory strategy by directing the student to read a sentence about circulatory system and then explain what new information each line provides and how it relates to what was previously read (see Chi, DeLeeuw, Chiu, & LaVancher, 1994). The term self-explanation is referred to explanations generated by the learner, which could be done by speaking aloud or in one's own head, written or typed. Providing feedback on the correctness of the explanation is beneficial. It is important to mention that generating incorrect self-explanations does not depress effective performance (Chi et al., 1989). Chi and her colleagues suggested that a possible mechanism underlying self-explanations of worked examples is producing a qualitative constraint network that represents knowledge of the solution steps, which possibly links together general theory and specific application. When a similar problem is encountered, qualitative propagation through the constrained network can yield a plan for a quantitative solution.

The effect of self-explanations can be explained from a cognitive load theory perspective. Learners who use cognitive and metacognitive elaboration strategies invest more mental effort, which is utilization of germane cognitive resources that stimulates construction of schemas. Renkl and Atkinson (2003) described the use of germane cognitive resources for different stages of skill acquisition: in early stages, germane resources are used to self-explain illustrated principles and generalize over presented worked examples. In later stages, when learners study worked examples in-depth, germane resources are utilized by anticipation of solution steps and imagining, and in the final stage, these resources are used for problem-solving. Renkl and Atkinson (2003) distinguished between the following self-explanation activities that have proven to be crucial: principle-based explanations (a learner assigns meaning to operators by identifying the principle), explication of goal-operator combinations (a learner assigns meaning to these operators by identifying sub-goals), and noticing coherence (a learner identifies connections among worked examples, which supports building abstract schemas).

Quality of self-explanations. Prior research identified considerable differences in learners' ability to self-explain (Chi et al., 1989; Renkl, 1997). In a study by Chi et al. (1989), the quality of self-explanations was measured by the number of inferences that fill in information gaps in the text. The term "high quality self-explanations" referred to generating inferences, integrating statements, and providing comments reflecting deep analyses of the text; and the term "low quality of self-explanations" referred as to paraphrasing and re-reading statements. Chi et al. found that learners who spontaneously generated a larger number of high quality self-explanations while studying incomplete worked examples scored significantly higher on post-tests than those learners that generated fewer high quality self-explanations. Renkl (1997) fixed the learning time for each individual in a study to isolate qualitative differences in self-explanation activities. He distinguished between successful and unsuccessful learners in the following main points: (1) principle-based explanations; (2) explication of goal-operator combinations; (3) anticipative reasoning; and (4) metacognitive monitoring. In addition, Renkl found that the successful learners frequently did not provide all of the types of self-explanations that were positively related to learning outcomes. According to Renkl, there are two types of successful learners: principle-based explainers and anticipated reasoners. Principle-based explainers are those who during their self-explanation activity mostly assign meaning to operators utilizing both principle-based explanations and explicating goal-operator combinations. Anticipated reasoners are those who mainly concentrate their effort on solution steps. Principle-based explainers did not frequently anticipate solution steps, in contrast to anticipative reasoners, who mainly concentrated their effort on solution steps and refrained from frequent principle-based explanations and explication of goal-operator combinations. Renkl (1997) identified two groups of unsuccessful learners: passive and superficial explainers. The passive explainers demonstrated a low level of self-explanation activity. Superficial explainers spend little time on studying worked examples. Renkl pointed out that most learners belong to the unsuccessful groups.

Prompts to self-explain. Self-explanation activity requires learners to invest mental effort into deep processing of information. Studies have shown that most learners do not spontaneously provide self-explanations while they study worked examples (Renkl, 1997; Renkl, et al., 1998). Renkl et al. (1998) suggested that a learning environment that combines the procedure with prompts to self-explain would encourage more active processing of worked examples. The authors suggested using prompts to elicit principle-based self-explanations at initial stages of learning, followed by procedures that induce anticipations to foster far transfer to improve schema formation. Similarly, Atkinson, Renkl, & Merril (2003) recommended the use of prompts to self-explain for teaching skills in complex subject domains because it enhances transfer performance and is relatively easy to implement without additional instructional time.

The effect of the described above cognitive load type-specific manipulations need to be empirically validated for instructional designers to be able to properly implement the proposed interventions into the design of instructions. However, research on cognitive load theory has not yet established type-specific measures of cognitive load (Ayers, 2006). In addition, this situation imposes challenges to testing CLT as a theory (Beckmann, 2010).

Measuring Cognitive Load

Recently, DeLeeuw and Mayer (2008) investigated separate measures for different types of cognitive load. DeLeeuw and Mayer suggested that: (a) response time to the secondary task is most sensitive to manipulations of extraneous processing created by adding redundant texts; (b) mental effort ratings during learning are sensitive to manipulations of intrinsic processing created by sentence complexity; and (c) difficulty ratings are most sensitive to differences related to germane processing reflected by transfer test performance. Manipulations of intrinsic cognitive load in DeLeeuw and Mayer's (2008) study were realized through the variations in complexity of the sentences learners had to process in order to perform the learning task. Beckmann (2010) who carefully analyzed DeLeeuw and Mayer's (2008) study pointed out that some sentences were sufficiently complex and adequately prepared learners for performing tasks, while other sentences were unnecessary complicated, and therefore posed additional extraneous load on learner's cognitive system. Beckmann concluded that it would be difficult to objectively differentiate between sufficiently complex sentences and unnecessarily complicated sentences. Beckman questioned whether the DeLeeuw and Mayer's manipulations actually varied sources of extraneous cognitive load, as opposed to intended intrinsic, and consequently considered assigned validity of effort ratings with regard to intrinsic load less than convincing.

Paas and his colleagues (Paas & van Merriënboer, 1993; Paas, Touvinen, Tabbers, & van Gerven, 2003) identified three main indicators of total cognitive load: mental effort, mental load, and performance. Mental effort is the cognitive capacity that is allocated to accommodate the demands imposed by the specific task (Paas et al, 2003). Mental effort could be measured by obtaining from the learner subjective ratings provided after the task completion. Mental load reflects total cognitive load imposed by a particular task on the learner's cognitive system. This load depends on task characteristics and the learner's level of expertise (Beckmann, 2010). The third main indicator of cognitive load is performance measures, such as posttest scores that directly measure learning outcomes.

The most common measurement of total cognitive load, developed by Paas (1992), is a 9-point rating scale for obtaining subjective measures of participant's perceived amount of invested mental effort. The 9-point mental effort rating scale can be used for multiple measurements during experiments, such as a single measurement after each task. This rating scale has demonstrated high internal consistency with reliability coefficient (Cronbach's alpha) in the range between .83 and .93 in several studies (Paas, van Merriënboer, & Adam, 1994; Paas, 1992; Paas & van Merriënboer, 1994; Kester, et al., 2004).

The reliance on subjective measures of total cognitive load alone will not provide conclusive results about the success in optimizing instructional design (Beckmann, 2010). Beckmann (2010) explained that each of the three possible outcomes of manipulations of cognitive load: decrease, no change, or increase could indicate both success and failure of an instructional intervention from a CLT perspective when only total cognitive load is measured. The decrease in total cognitive load does not provide enough information to conclude whether the intervention was successful. It signifies that extraneous load was reduced; however, does not clarify whether cognitive resources were redirected into germane activities. No change either suggests that none of cognitive load types were affected by design manipulations, or that extraneous cognitive load was reduced and the use of germane resources increased. According to Beckmann, an increase in in mental effort ratings indicates a partial success and suggests higher levels of germane activity, while the design manipulation failed to reduce sources of extraneous cognitive load. One could argue with this statement and suggest that an increase in total load would rather indicate that the intervention did more harm than good from a CLT-perspective. For

CLT-based interventions, reference to objective performance measures combined with subjective ratings could help researchers to decide whether an intervention was successful (Beckmann, 2010). Based on the assumption that cognitive load manipulations enable learners to process information faster and easier, performance measures combined with time on task or mental effort invested in the task completion can reveal important information about cognitive load (Beckmann, 2010). Learning efficiency, suggested by Morrison, Ross, and O'Dell (1988) measured level of achievement attained per allocated instructional time for CBI instructions. Efficiency scores were computed as a ratio of a total posttest score and lesson completion time. Instructional efficiency measure developed by Paas and van Merriënboer (1993) as a combination of test performance measures and intensity of mental effort invested into task completion is another measure that can provide a good estimator of cognitive load and consequently of the effectiveness of an instructional intervention (Sweller et al., 1998; Beckmann, 2010).

Efficiency of Instructional Condition. Efficiency of instructional condition measure combines measures of test performance with measures of mental effort invested to attain this test performance (Paas and van Merriënboer, 1993). This measure can be calculated based on two separately obtained measures: participants' post-test performance and 9-point rating scale measures of perceived amount of invested mental effort. First, measures of the invested mental effort and performance have to be standardized (the mean value has to be subtracted from each participant's value, and the result divided by standard deviation), yielding in each participant's z-score for the invested mental effort and a z-score for performance. Second, instructional condition efficiency score can be computed for each participant using the following formula:

$$E = \frac{z_{performance} - z_{mental\_effort}}{\sqrt{2}}$$

This formula was derived using the perpendicular distance of a point defined by two coordinates in Cartesian coordinate system with invested mental effort on x-axis and performance on y-axis to a zero-efficiency line, where mental effort is equal to performance. Mental efficiency can be visualized using graphical representations. When considering these graphs, one should note that the instructional conditions located above the diagonal line (mental effort = performance) have greater relative efficiency scores because they have a higher group performance with lower invested mental effort (Paas & van Merriënboer, 1993). Conversely, low instructional efficiency (below the line) is the result of low task performance and high mental effort invested (Tuovinen & Paas, 2004). Figure 1 represents relative instructional condition efficiency for the three conditions: conventional problem-solving, worked examples, and completion problems, obtained based in the results of the experiment presented in Table 2, as adapted from Paas and van Merriënboer (1993).

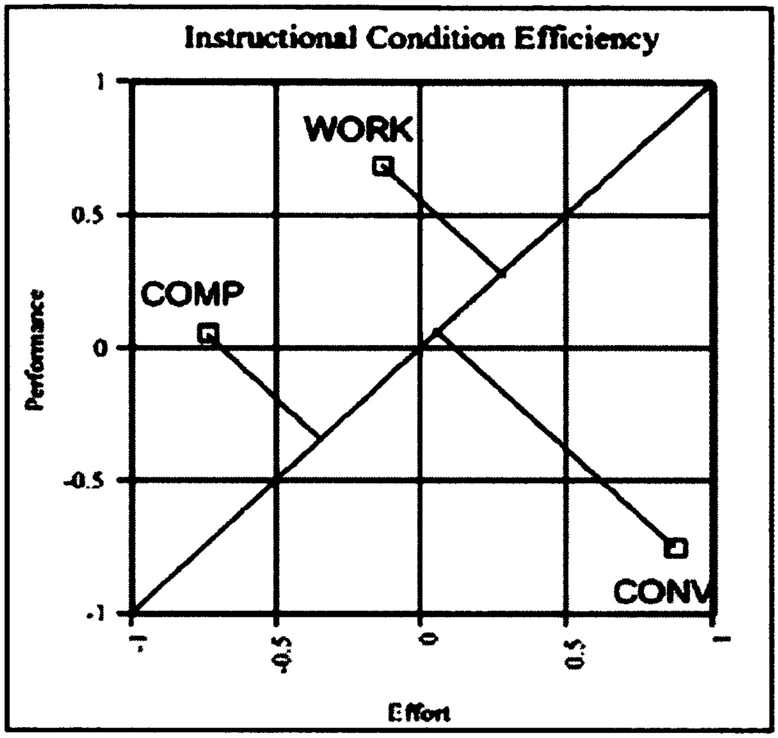


Figure 1. Instructional Condition Efficiency. Adapted from "The efficiency of instructional conditions: An approach to combine mental effort and performance measures," by F. Paas, and J. J. G. van Merriënboer, 1993, *Human Factors, 35*(4), 737–743.

Table 1

Effort-Performance and Relative Condition Efficiency Means (Paas and van Merriënboer, 1993)

Instructional condition	Effort	Performance	Relative condition efficiency
Conventional problems	0.87	-0.75	-1.15
Worked examples	-0.13	0.69	0.58
Completion problems	-0.74	0.06	0.57

Note. Adapted from "The efficiency of instructional conditions: An approach to combine mental effort and performance measures," by F. Paas, and J. J. G. van Merriënboer, 1993, *Human Factors, 35*(4), 737–743.

The perpendicular distance from the neutral efficiency condition (diagonal line) where effort equals to performance to each of the points plotted in the effort-performance coordinate system, according to the group effort and performance means, represents the efficiency values for each group. As shown on the graph, the conventional problem-solving group demonstrated the lowest efficiency, as compared to worked examples and completion problems that have demonstrated almost equal efficiency.

Van Gog and Paas (2008) revisited the formula for the efficiency of instructional condition that was adapted with some modifications to the original one and strongly recommended studies to use the original formula. The authors gave two important recommendations to researchers analyzing efficiency of an instructional condition that decrease extraneous load and increase germane activity: (1) use the 9-point rating scale based on perceived amount of invested mental effort, rather than a 7-point perceived task difficulty rating scale; and (2) use mental effort rating scale and performance scores during the testing phase, rather than mental effort during the learning phase and performance scores for the testing phase. Van Gog and Paas analyzed commonly used practices in obtaining subjective mental effort measures and concluded that the 9-point mental effort rating is a more precise instrument for measuring total cognitive load than widely adopted 7-point rating scale (e.g., Kalyuga, Chandler, & Sweller, 1998, 1999, 2000, 2001; Moreno & Valdez, 2005, as sited in van Gog & Paas, 2008) asking participants not to rate mental effort, but to rate how difficult they perceived the task. The authors explained that although the concepts of invested mental effort and perceived task difficulty are related, asking students to rate how much mental effort they invested in completing the task versus how difficult they perceived a task are two different questions that can lead to different interpretations. According to van Gog and Pass, the first question (invested mental effort) pertains to a process, and the perception will likely involve more aspects than only the task itself, whereas the second question (perceived task difficulty) pertains mainly to the task. Therefore, questions of perceived task difficulty in comparison to invested mental effort will most likely lead to nonequivalent ratings, presumably on the extreme end of the scale. For example, Paas, Tuovien, van Merriënboer, and Darabi (2005) have shown that when learners perceive a problem to be extremely difficult, they may not be motivated to invest much effort into it.

Van Gog and Paas (2008) explained that measures of mental effort in combination with performance measures during the task performance provide researches with a better indicator of the quality of learning outcomes. For example, in case of the study that aims to reduce extraneous and intrinsic cognitive load and increase the use of germane resources, it is expected that learners in the treatment condition would require more mental effort investment during the learning phase. One could expect that because this effort was invested in learning, higher performance score and a less mental effort on the test compared to non-treatment group. However, if the invested mental effort was only measured during the learning phase it could be higher than during the actual test, and therefore efficiency of instructional condition will result in incorrect lower numeric value. Although efficiency of instructional condition was not measured in Sliva, Morrison, and Watson (2011) study, the presented argument suggests why no significant difference in cognitive load measures taken during the learning phase were found, although the study demonstrated higher transfer test performance for the experimental condition where participants were prompted to self-explain and plant, compared to non-prompted (control) condition. Although total cognitive load remained the same for both groups the control and experimental, better learning outcomes in the experimental group suggest the following explanation. Since intrinsic load was held constant for the learners with the same level of expertise, the change in total load remained constant. Since better learning outcomes occurred, the results suggest that during the testing phase extraneous load was further reduced and the use of germane resources was increased for the experimental group, because the use of extraneous and germane resources are defined to have opposite effects on learning: detrimental or beneficial, respectively (Ayres, 2006; van Gog & Paas, 2008).

The use of relative mental efficiency of instructional condition allows the learner to combine the effort invested during performance with the actual performance on a test. This effect, in turn, allows the researcher to compare instructional formats not only in terms of their effectiveness but also in terms of their efficiency. A number of studies have demonstrated the added value of the efficiency measure by showing that differences in the effectiveness of instructional condition are not always identical to differences in their efficiency (Paas & van Merriënboer, 1994; Pollock, et al., 2002; van Gerven, Paas, van Merriënboer, & Schmidt, 2002; van Gerven, Paas, van Merriënboer, Hendriks, & Schmidt, 2003; van Merriënboer, et al., 2002).

Purpose of the Research

The purpose of this study was to compare the use of prompts to self-explain with and without completion problems, and conventional problem-solving practice.

Research Questions and Hypotheses

The study tested three hypotheses and asked two research questions. The following were the three hypotheses:

1. Perceived cognitive load reported by participants in two experimental conditions (i.e., prompted to self-explain and prompted to self-explain completion problems) will be

lower than perceived cognitive load reported by participants in the control condition during both learning and testing phases.

2. Participants in the two experimental conditions will perform better than participants in the control condition on quizzes and a posttest.
3. The two experimental conditions will have higher efficiency than the control condition.

The study also aimed to answer the following research questions:

1. Will the quality of self-explanations change throughout the duration of the study for the students in two experimental conditions?
2. Will the treatment groups show a more favorable attitude towards the instructional strategy?

CHAPTER II

METHODS

Participants

Initially, 63 undergraduate students enrolled in a semester-long calculus-based introductory physics course (that is, one lecture section) at an urban university in the United States consented to participate in this study. Participants were randomly assigned to three equal groups (i.e., self-explanation, self-explanation & completion, and control groups) of 21 students each. However, data from only those participants who completed all the work required for this study was taken in consideration resulting in adjusted self-explanation (n = 19), adjusted self-explanation & completion (n=18), and control (n = 19) groups; a total of 56 participants. Participants were majors in the following fields: computer science, chemistry, biology, mathematics, physics, exercise science, and engineering. There were no direct benefits offered for participating in this study. When students were invited to participate in this study, they were offered an opportunity to be introduced to instructional techniques that could improve their problem-solving skills.

Research Design

The study utilized 3 (between) x 4 (within) fixed-effects design. Treatments were instructional condition (self-explanation, self-explanation & completion, and control) measured weekly for the four weeks. A two-way ANOVA was used to compare three treatment conditions measured weekly over the four week study period for the following measures: cognitive load for homework assignments and quizzes; time on task for homework assignments; and performance for quizzes. A two-way ANOVA was used to compare two prompted self-explain conditions (i.e., between subjects) measured weekly over the four-week study period (i.e., within subjects) for the quality of self-explanations on homework assignments. Further ANCOVA was used to compare how quality of self-explanations changed between two experimental groups with time (weekly submission number used a covariate). In addition, one-way ANOVA was used to compare the quality of self-explanations between the first and the last submission for each experimental group separately. The study also utilized a one-way ANOVA to compare three treatment conditions (between subjects) as measured for one posttest for cognitive load and performance. Since efficiency score was based on normalized measures of performance and mental effort, a simple comparison for average values of efficiency was performed between groups. And finally, between groups comparison of eight survey items requiring responses on a 7-point rating scale was performed using one-way ANOVA; comments to the remaining six survey items were analyzed qualitatively.

In-Class Instruction. All students participating in this study received the same in-class instruction throughout the study. This strategy ensured that content was the same across all treatments. All students attended the same lecture together. During the initial in-class information presentation, extraneous and intrinsic cognitive load were reduced for all treatment groups in the following manner.

To reduce intrinsic cognitive load, this study utilized presentation of information in two phases: first, the element interactivity of materials was reduced by presenting interacting elements in isolation, and second, all the information was presented in its full-complexity (Pollock et al., 2002). And a simple-to-complex sequence with general supportive information presented first, before an equivalent learning task (van Merriënboer et al., 2003) was used.

Extraneous cognitive load during the lectures was reduced by using integrated text and diagram formats in the PowerPoint slides, avoiding presentation of redundant information, and using audio and visual modalities to present mutually referring textual and pictorial information (i.e., the instructor also drew diagrams and simultaneously verbally explained the interrelations between the elements in these diagrams), as suggested by Sweller et al. (1998). In addition, to support acquisition of problem-solving skills, worked examples were integrated into class presentations (Sweller, 1999). Unique worked examples were designed by the instructor and were different from assigned homework problems.

Treatments

All participants were assigned the same seven homework problems each week. The following paragraphs explain how the treatments differed.

Control Treatment. The control group was simply instructed to solve homework problems and did not receive any explicit prompts or partial solution. For example, for a unit on the conservation of energy, students were asked to solve the following problem at home: A spring of constant k=340 N/m is used to launch at 1.5-kg block along a horizontal surface whose coefficient of sliding friction is 0.27. If the spring is compressed 18 cm, how far does the block slide?

Self-Explanation Treatment. Participants in this group received the following prompts *to self-explain*: (1) name the topic you are investigating; (2) state the general principle that can be used to solve this problem; (3) write down the formula for that principle; (4) adjust this formula for the specific situation and name each component of the obtained formula; (5) using the results of the previous step, derive an equation or a system of equations for solving the problem; (6) solve that equation or a system of equations for the unknown variable, and find a numeric answer; and (7) check your answer in terms of measurement units and from a common sense point of view. For example, the final prompt for checking answer in terms of measurement units was, "To check the measurement units means to confirm that the units derived from a final algebraic solution correspond to the measurement units, in which to-be-found value is measured." According to the instructor of the course, students do not check measurement units on the regular basis, if they are not specifically instructed to do so. Participants in the self-explain condition received initial training in how to self-explain. This training consisted of a PowerPoint tutorial in which each of the steps of self-explanation activity was outlined and helpful hints on how to better self-explain were provided. This tutorial can be found in Appendix A.

Self-Explanation & Completion Treatment. Participants in this group were given partial solutions to the problems in addition to the prompts to self-explain. Completion problems were posted weekly in a special section of Blackboard for the prompted self-explain completion group. Other treatment groups could not view these problems. A partial solution was designed to help students in identifying the goal state of the problem. Such help could be in the form of formula(s) for the principle for solving the problem and formula(s) for the application of that principle to a specific task (i.e., operands to use in algebraic form). All other intermediate solution steps and the final answer were omitted requiring learners to complete the solution. For example, self-explanation & completion group was assigned the same problem on conservation of energy. Participants in this group could see the following partial solution on Blackboard:

$$\Delta K + \Delta U = W_{\mathrm{nc}}$$

$$U_e = \frac{1}{2}kx^2$$

$$W_{\mathrm{nc}} = -\mu_k mgL$$

Participants in this group were also directed to follow the same seven-step prompts as asked of the self-explanation group and received the same initial training on how to self-explain.

Example handouts for each group can be found in Appendix B. Sample homework submissions for each of the groups can be found in Appendix C.

During the semester, all students enrolled in a calculus-based introductory Physics course were required to do their homework online. Students were presented with the problem in multiple choice question format with answers in a numeric form. Students were directed to solve the problem with pencil and paper and then select an answer that best represented their solution from the answer options. Since students had multiple attempts to find the right answer, such practice often resulted in guessing instead of solving the problem. Although the instructor required a specific format for solving problems, he had no control over students' problem-solving practice at home. During this study, participants were required to write down their solutions on paper and submit them to the researcher, as opposed to simply selecting the answer, which could result in better quality solutions, regardless of the treatment group. However, participants prompted to self-explain were presented with the general systematic approach on how to solve physics problems. Participants prompted to self-explain completion problems were provided with both a partial support in identifying the goal state and a systematic approach to solving problems. In contrast, the control group might be more likely to implement a means-ends analysis (Sweller, 1988).

Instruments

The following is the description of instruments that were used to measure quiz- and post-test performance, cognitive load, efficiency, quality of self-explanations, and learners' attitude towards the instruction.

Quizzes. Four weekly quizzes were administered to all participants during this study. At the start of the week following the instructional presentation and submission of homework, all students were given a closed-book and closed-notes quiz based on the materials presented during the previous week. Students had 20 minutes to complete the quiz.

Quiz design. This quiz consisted of one problem with five sub-questions each requiring free response. From these five sub-questions, one measured comprehension, two sub-questions measured near transfer, and two sub-questions measured far transfer.

Quiz grading. Quizzes were graded by the course instructor. A total of 22 maximum points could be awarded for the quiz that is two possible points with no partial credit for the sub-question measuring comprehension, and five possible points using partial credit for each near and far transfer sub-question. The rubric that was used to assign partial credit (as adapted from the course instructor) can be found in Appendix D. Quizzes [1-4] produced Cronbach's Alpha coefficient varying between 0.72 and 0.74 (i.e., 0.72, 0.74, 0.72, and 0.73, respectively).

Posttest. At the end of the fifth week counted from the day the study began, and after the fourth quiz was completed and feedback provided, all participants completed a posttest based on eleven instructional objectives covered during the four preceding weeks of the study. This test was adapted directly from the instructor's practices.

Posttest design. The posttest consisted of 11 problems that included four rote level problems testing recall; four problems similar to those used for weekly quizzes testing near and far transfer of performance, as well as comprehension; and three story problems specifically testing far transfer, as well as learner's ability to translate a word problem into physics context. Rote level problems were presented in a multiple-choice format, and the remainder of the problems required free responses. Some of the problems consisted of several scored elements resulting with a total of 27 items on this posttest. Students had 1 hour and 50 minutes to complete this test.

Posttest grading. A posttest was graded by the course instructor. Rote level problems were worth one point each without a partial credit option for a maximum of four points possible. Partial credit for the four near and far transfer problems was awarded in the same manner as for the quiz problems (see quiz grading). Each story problem was worth a maximum of 10 points. The rubric for the score awarded for the story problem can be found in Appendix E. And the summary of the posttest grading can be found in Appendix F.

The maximum score for the posttest was 122. Scoring consisted of a maximum of four points possible for recall, eight points possible for comprehension; a maximum of 40 points possible for the near transfer; and a maximum of 70 points possible for the far transfer. A posttest grade was calculated as a percentage of total points awarded out of maximum possible points. Cronbach's Alpha for the test was 0.84.

Cognitive Load Measures. Cognitive load measures were taken after completion of each homework problem, each quiz, and for after solving each posttest problem resulting in 43 measures of cognitive load for each participant. Participants in each treatment condition were instructed to rate their invested mental effort on a 9-point Likert-type scale, as adopted from Paas (1992). To rate invested mental effort, students were asked: “How much mental effort did you invest in solving this problem?" and rate their responses from1 (very, very low) to 9 (very, very high) on a 9-point scale. Example of the scale is provided below:

1--------2--------3--------4--------5--------6--------7--------8--------9
very, very,
very low very high

Figure 2. Likert Scale for Invested Mental Effort.

The following table explains how many cognitive load measures were recorded for each participant on each task. Each recording was treated as an individual score. Cognitive load measures were calculated as mean values.

Table 2

Measures of Cognitive Load

Task	Number	When Measured	Total Recordings	Total
Homework	7	After each problem	4	28
Quizzes	1	After completing quiz	4	4
Posttest	11	After each problem	1	11

A total of 63 students participated in this study. However, only data from 56 participants were used for the purpose of this study because data from seven participants were incomplete or missing. The following table explains how many participants failed to provide cognitive load measures, the reason for that failure, and a total number of participants whose data were collected and used in this study.

Table 3

Measures of Cognitive Load Recorded for Each Participant

Task	Number of participants	Number of participants failed to accomplish	Reason	Number of participants reported in the study
Homework		1	Reported 1 CL measure on submissions 1 and 2	
Quizzes		5	Not all 4 quizzes taken	
Posttest		1	No CL measures reported	
Total	63	7		56

A total of 43 cognitive load measures were recorded for each of the 56 participants on a 9-point Likert scale, resulting in a total of 2,408 recordings. Cronbach's Alpha for the rating scale combining all cognitive load measures was 0.89.

Additional Data: Efficiency Measures

The efficiency of instructional condition scores (as adopted from Paas & van Merriënboer, 1993) were calculated based on measures of participants' performance on the quizzes, posttest, and self-reported measures of mental effort. For the four quizzes, each student's score on each quiz was combined with self-reported measure of cognitive load for the quiz. For the posttest, each student's score on each posttest problem was combined with the self-reported measure of cognitive load for that problem. Efficiency of instructional condition was calculated in the following way. First, both student's performance measure and invested into that performance mental effort measure was standardized, yielding in each participant's z-score for performance and a z-score for the invested mental effort. Second, instructional condition efficiency score was computed for each student, using the following formula:

$$E = \frac{z_{performance} - z_{mental\_effort}}{\sqrt{2}}$$

For each student, one efficiency measure was calculated for each quiz (a total of four quizzes) and 11 efficiency measures were calculated on a posttest.

Time on task. Since time on task was limited for quizzes (i.e., 20 minutes) and a posttest (i.e., 1 hour and 50 minutes), time on task measures were only obtained for homework submissions. After completion of each homework problem, participants in all four treatment groups self-reported the time they spent solving the problem. They were asked: "How much time did it take you to solve the problem?" and recorded time in hours (if applicable) and minutes. For each student, seven time measures (minutes spent on each problem) were recorded for each homework submission (a total of four weekly submissions).

Quality of Self-Explanations Measures. To obtain a score on the quality of self-explanations, two independent raters evaluated each of the seven homework problems submitted across the four homework assignments by the participants in the self-explanation and self-explanation & completion groups. A rating of zero to three was assigned to each of the seven prescribed self-explanation steps ranging from no points when no attempt was made to complete that step to points of one, two and three to represent a failed attempt, a partially correct attempt, and a completely correct attempt, respectively. Points from the seven steps were added together for a total score of 0 to 21, adapted from Sliva et al. (2011) study. This rubric can be found in Appendix G.

Two raters were recruited from TAs working with the course instructor during the semester. They received initial instructions on how to evaluate homework submissions for the quality of self-explanations on the first week of the study. The researcher was available to meet with raters during the study to discuss any questions or concerns. Inter-rater reliability was assessed using Cronbach's Alpha coefficient. Cronbach's Alpha estimated the correlation between two independent raters and allowed to interpret the result as an estimate of interrater reliability. For this study, Cronbach's alpha coefficient of 0.96 (alpha>0.90) suggested excellent agreement between raters. One quality of self-explanation score was assigned for each student on each submission. This score was calculated as a median value of all quality scores measured by both raters on all problems for that submission. Cronbach's Alpha for the instrument measuring quality of-self-explanation score in this study was also 0.96.

Attitude Survey. A 15-item attitude survey was administered immediately following the posttest in order to better understand the participants' attitude toward the instruction. In the survey, participants were asked to report their treatment group and then asked to rate their overall agreement on a 7-point Likert-type scale ranging from "strongly agree" to "strongly disagree" for eight statements related to the effectiveness of the treatment, changes that it may have induced, and the likelihood that the participant would continue using this method in the future. The scores for these Likert items were calculated as means. Cronbach's Alpha for these eight items was 0.93. In addition, participants completed six open-ended response items regarding what they especially liked or disliked about the study and any other thoughts or comments about their experiences with the instructional method. Open responses were interpreted qualitatively. A copy of the survey is located in Appendix H.

Procedure

The research was conducted during a six-week period in the fall semester of 2012, starting approximately three weeks after the start of the semester. Students met with their instructor twice a week for a 1 hour and 50 minute lecture. Approximately one week prior to the study, the researcher introduced the study to the class and distributed and collected Informed Consent Forms. The copy of the Informed Consent Form can be found in Appendix I. Consenting participants then were randomly assigned to one of three treatment groups. Participants were not explicitly told to which treatment group they were assigned and groups were coded as A, B, and C to minimize their knowledge of the treatments.

Participants in the two self-explain treatments received training on self-explanations during the last 15 minutes of the lecture prior to the beginning of the study. PowerPoint tutorial on how to self-explain used for this training was available to participants in the two experimental groups on BlackBoard throughout the duration of the study. For the first homework assignment consisting of seven problems, participants in all three treatment groups received handouts with detailed instructions on how to prepare and submit homework assignments. In addition, examples on how each group had to submit their homework assignments were posted online separately for each treatment group. Participants were required to review self-explanation tutorial prior to submitting the first homework assignment. Participants were also required to follow the instructions in their handouts while working on their first homework assignment. During subsequent study weeks, handouts and examples were available separately online to each treatment group and participants were asked to review these when working on homework assignments.

The same in-class instructions were used for all the treatment groups with all participants attending the same lecture (i.e., at the same time), thus there was no variability in the lecture instruction. The instructor delivered instructional materials and demonstrated worked examples in the same manner for all students during the lectures. The researcher was present at all lectures to record any variations in the amount of explanations provided by the instructor in class to ensure that lectures were not designed in favor of self-explanation technique.

During the lectures, the instructor introduced facts, concepts, and principles to all students together in the same auditorium. After a brief introduction of historical facts, the instructor presented students with formal definitions of the concepts, pointed out their critical attributes, and compared and contrasted these attributes for closely related concepts. During the lecture, the instructor used the RULEG approach. When using RULEG, the instruction begins with the statement of the rule or principle that explains the relationships between concepts that is further followed by the application of this rule or principle (Morrison, Ross, Kalman, & Kemp, 2011). In this study, application of the rule or principle consisted of multiple worked examples. The number of worked examples demonstrated in class varied dependent on topic complexity and available time. These worked examples were presented to all students using the format the instructor requires for solving physics problems in his class: (1) state what is given and what needs to be found, (2) create and label a related diagram, and (3) solve the problem step-by-step. It is important to note that explanations provided by the instructor during his lectures were fundamentally different from the self-explanation activity because self-explanations were generated by the learner, as contrasted with explanations provided by any external source (Chi et al., 1989). Careful attention was given to ensure that the instructor did not demonstrate or model the prompting or self-explanation activities during the class presentation.

Each week, participants in all groups were assigned an average of seven homework problems to solve using pencil and paper outside of class. These problems were different from unique worked examples designed by the instructor to introduce in class topics included in the course syllabus. These problems were the same for all three groups and were adapted from the textbook selected by the university physics department for this class. Homework problems were related to the same principles discussed in class; they contained the same variables but in different contexts. The researcher worked closely with the instructor on the selection of homework problems and their modification into completion problems. This process ensured that problems adequately represented all topics discussed in class in terms of difficulty levels, that is, problems include comprehension, and that they measured both near and far transfer of performance. After the instructor selected problems for the weekly homework assignment, the instructor and the researcher worked together to modify assigned problems to completion problems using the solution manual that accompanied the textbook. Intermediate steps and the final answer in the complete solutions were omitted leaving only partial solutions for the students to complete. Homework problems were posted weekly on Blackboard. Problems, as well as partial solutions, and prompts to self-explain were posted online separately for the treatment groups at the beginning of the week. Participants received weekly notifications through university email from their instructor that the assignments were ready.

At the end of the week, homework was collected by the researcher in class and copied over the weekend for further analysis. It was returned back to students at the beginning of the following week.

For the two experimental groups, the quality of self-explanations was evaluated according to the rubric provided in the method section, and quality of self-explanations scores were recorded but not disclosed to the participants. Those scores were used only for the study and were not used to calculate the course grade. Participants who did not follow the prompts or provided poor self-explanations were notified by the researcher that they needed improvement. Feedback was provided by writing notes in the students' journals with individual instructions. Two additional training sessions were held based on a PowerPoint tutorial and took into account students' common misconceptions.

Weekly quizzes were administered under the instructor's supervision on Tuesdays and lasted approximately 20 minutes. One post-test that lasted 1 hour and 50 minutes was administered at the end of the study to all participants. Quizzes and the posttest had to be completed using pencil and paper. The researcher was present at all testing events to ensure that participants did not forget to rate the invested mental effort for each problem they solved.

At the end of the session immediately following the posttest, participants responded to a short attitude survey that took approximately 15 minutes to complete.

Data

Dependent measures, details on data collection, and analyses for the three hypotheses and one research question are presented in the following table.

Table 4

Data Related to Research Questions and Hypotheses

Hypotheses/ Research question	Dependent variable	Timing	Instrument	Comparison
Hypothesis 1	Self-reported CL measures for each student: a) homework problems (four submissions with seven problems each) b) quiz problems (four quizzes with one problem each) c) on each of the eleven post-test problems	After completion of *each* individual homework problem and post-test problem. After completion of each quiz.	9-point Likert-type rating scale for the invested into solving the problem mental effort (Paas, 1992)	Two-way ANOVA (between subjects-group; within subjects-week) for weekly homework assignments and quizzes; One-way ANOVA (between subjects-group) for the post-test
Hypothesis 2	a) Each student's quiz grade, and near and far transfer score b) Each student's post-test grade, and near and far transfer score	Each time testing is administered	Adapted from instructor's rubric for grading	Two-way ANOVA (between subjects-group; within subjects-week) for weekly quizzes; One-way ANOVA (between subjects-group) for the post-test

Hypotheses/ Research question	Dependent variable	Timing	Instrument	Comparison
Hypothesis 3	a) Efficiency score for each student on each quiz and a posttest problem	a) Since obtaining efficiency score doesn't require separate measures, it can be done during data analysis process, after the study is over	a) Formula for calculating the efficiency of instructional condition (Paas & van Merriënboer, 1993)	Comparison of means for efficiency formula, comprised of z-scores of performance and z-scores of mental effort and their graphical representation using the perpendicular distance from the line performance=effort
	b) Time on task for each student on each homework problem (four submissions with seven problems each)	b) After completion of each problem	b) Self-reported by each participant amount of time used for solving the problem	Two-way ANOVA (between subjects-group; within subjects-week) for weekly homework assignments
Research Question 1	Quality of self-explanation for each student in two prompted self-explain (with and without partial solution) conditions for each homework problem	After the study is over	Rubric for the quality of self-explanations, as adapted from Sliva et al. (2011) study, done by two independent raters	Two-way ANOVA (between subjects-group; within subjects-week) for weekly homework assignments ANCOVA (group-two levels as a fixed factor, submission-four levels as covariate) for four weekly homework assignments One-way ANOVA (Welch F-ratio test) for each group separately, between the first and the last assignment
Research Question 2	Response to 8 survey items	At the end of the study	7-point rating scale from “strongly agree” to “strongly disagree”	One-way ANOVA (between subjects-group) for each survey item separately

In addition, to answer the second research question, responses to the attitude survey 6 items that required comments were analyzed qualitatively.

CHAPTER III

RESULTS

This study aimed to test three hypothesis and answer two research questions. The reported results for the hypotheses and the first research question were obtained using quantitative analyses, and the second research question was answered using quantitative and qualitative analyses.

Hypothesis I

The *first hypothesis* predicted a lower cognitive load for the two experimental conditions compared to the control in both learning and testing phase. In this study, total cognitive load was obtained using the most common measurement, developed by Paas (1992), using a 9-point rating scale of participant's perceived amount of invested mental effort (see method section).

One-way between subjects ANOVA was used to compare means of subjective cognitive load for the posttest. The overall effect of group was significant, $F(2,613)=11.25$, $p<.001$, partial $\eta^2 = .06$. Since homogeneity of variances was violated (Levene's test, $F(2,613)=0.71$, $p=.49$), Bonferroni test was used for a post-hoc comparison, resulting in self-explanation group reporting significantly higher cognitive load ($M=6.70$, $SD=2.35$) than the control group ($M=5.33$, $SD=2.51$) and self-explanation & completion group ($M=5.66$, $SD=2.38$); there was no significant difference found between the control and self-explanation & completion group.

Two-way ANOVA (between subjects-group, within subjects-weekly quiz number) was used to compare means for subjective cognitive load measures reported by participants on four weekly quizzes. The main effect $F(2,212)=10.68$, $p<.001$, $\eta^2 = .03$ was significant. Since Levene's test ($F(11, 212)=1.59$, $p=0.09$) did not reject null hypothesis, Bonferroni test was used for a post-hoc comparison resulting in self-explanation group (M=6.62, SD=1.89) and self-explanation & completion group (M=6.42, SD=1.89) both reporting significantly higher cognitive load than the control group (M=5.85, SD=2.10) with no significant difference found between these two groups. The interaction effect was not significant, F(6,212)=2.21, p=.60.

For the homework, two-way ANOVA (between subjects-group, within subjects-weekly submission number) was used to compare cognitive load reported by participants after each problem. In this analysis, submission number [1-4] represents time, since homework was collected weekly within a period of 4 weeks. For example, submission 2 represents homework collected at the end of the second week. The main effect was significant, F(2, 1612)=35.45, p<.001, $\eta^2 = .04$. Since homogeneity of variances condition was met, Levine's test, F(11, 1612)= 4.06, p<.001, Turkey HSD test was used for a post-hoc comparison resulting in significantly higher reported cognitive load for self-explanation group (M=6.04, SD=2.13) and self-explanation & completion group (M=5.67, SD=1.96) as compared to the control (M=4.97, SD=2.36), with self-explanation group reporting significantly higher cognitive load than self-explanation & completion. The interaction effect was not significant, F(6, 1612)=1.58, p=.15.

In summary, self-explanation group reported a significantly higher cognitive load compared to the control on the posttest; both experimental groups reported a significantly higher cognitive load than the control group on quizzes and homework. In addition for the homework, self-explanation group reported a significantly higher cognitive load than self-explanation & completion group. Although a significant difference in cognitive load was found between two experimental and the control conditions during both learning and testing phase, the reported cognitive load in two experimental conditions was not different than the direction predicted. That is, the treatments produced a higher rating of cognitive load rather than reducing cognitive load. The first hypothesis was not supported.

Hypothesis II

The *second hypothesis* predicted better posttest and quiz performance for the two experimental conditions compared to the control.

For the posttest, one-way ANOVA was used to compare grades between groups. The group effect was significant F(2,53)=3.30, p<.05, $\eta^2 = .11$. Since Levene's test (F(2,53)=2.19, p=.12) did not allow to reject null hypothesis, Bonferroni test was used for a post-hoc comparison resulting in self-explanation group (M=66.37, SD=11.03) significantly outperforming the control group (M=56.78, SD=13.73); self-explanation & completion group (M=63,22, SD=14.26) was not significantly different from the control or from the self-explanation group.

Separate analysis of near transfer performance for the posttest revealed a significant effect of group, F(2,53)=7.74, p=.001, $\eta^2 = .23$. Since homogeneity of variances was violated (Levene's test, F(2,53)=1.24, p=.30), Bonferroni test was used for a post-hoc comparison. Both self-explanation (M=28.37, SD=3.74) and self-explanation & completion (M=27.56, SD=4.73) groups significantly outperformed the control (M=23.05, SD=4.88). No significant difference was found between the two experimental groups on near transfer of performance. Separate analysis of far transfer of performance for the posttest resulted in a significant effect for group F(2,53)=3.28, p<.05, $\eta^2 = .11$. Since homogeneity of variances was violated (Levene's test, F(2,53)=1.42, p=0.25)), post-hoc Bonferroni test revealed that self-explanation group (M=49.95, SD=5.78) significantly outperformed the control (M=43.68, SD =8.76) but self-explanation & completion group (M=46.17, SD=7.92) was not significantly different from the control or from the self-explanation group.

For the quizzes, two-way ANOVA (between subjects-group, within subjects-quiz number) was utilized to compare means of performance. The main effect was found to be significant, F(2, 212)=3.05, $p<.05$, $\eta^2 = .03$. Since homogeneity of variances was assumed, Turkey HSD test revealed a near significant difference (p=0.48) between the control group (M=69.25, SD=26.37) and self-explanation group (M=78.03, SD=26.37); self-explanation & completion group (M=69.46, SE=2.92) was not significantly different from the control or self-explanation group. The interaction effect was not significant, F(6, 212)=0.693, p=.66.

A separate analyses for near transfer quiz performance utilizing two-way ANOVA resulted in non-significant main effect, F (2,212)=2.12, p=0.12, that is the control group (M=7.57, SD=1.67), self-explanation group (M=8.01, SD=1.39), and self-explanation & completion group (M=7.56, SD=1.60) did not perform differently on quiz questions testing near transfer of performance. The interaction effect was also not significant, F(6, 212)=0.693, p=.66, $\eta^2 = .02$. As for far transfer, the effect of group was significant, F(2,212)=4.49, $p<0.05$, $\eta^2 = .04$. Since Levene's test (F(11,212)=1.62, p=0.10) failed to reject the null hypothesis, Bonferroni post-hoc test was used for further comparison. It revealed that the self-explanation group (M=8.63, SD=4.75) significantly outperformed both the control (M=6.68, SD=4.61) and self-explanation & completion group (M=6.71, SD=4.34), with no significant difference between the self-explanation & completion and control groups. The interaction effect was not significant, F(6, 212)=1.10, p=.36.

Overall, the second hypothesis was partially supported. Compared to the control group, self-explanation group demonstrated a better performance on quizzes and the posttest, and self-explanation & completion group demonstrated better performance on near transfer posttest questions only.

Hypothesis III

The *third hypothesis* predicted that two experimental conditions would have higher efficiency than the control condition. Two measures were used to analyze efficiency: time on task for homework problems and instructional condition efficiency for each of the four quizzes and the posttest, calculated as mean values for z-scores of performance and mental effort (see method section) for the three treatment groups.

Time on task for homework problems was compared with two-way ANOVA (between subjects-group, within subjects-weekly submission number). The main effect was not significant F(2, 1612)=1.96, p=.14, $\eta^2 = .002$. The control group (M=14.98, SD=15.56), self-explanation group (M=13.95, SD=12.03), and self-explanation & completion group (M=15.54, SD=11.20) spent on average the same time for each homework problem (time was measured in minutes). The interaction (treatment by homework submission) effect was significant, F(6, 1612)=2.50, p<.05, $\eta^2 = .01$. Since homogeneity of variances was assumed, (Levene's test, F(11, 1612)=3.55, p<.001), Turkey HSD test was used for a post-hoc comparison that revealed a significantly higher time on task for self-explanation & completion group (M=15.29, SD=10.32) as compared to self-explanation group (M=12.82, SD=9.57) on the second week homework submission, and a significantly lower time on task for the control group (M=13.47, SD=8.74) as compared to self-explanation & completion group (M=17.67, SD=14.51) on the fourth homework submission (see Figure 1). All other interaction effects were not significant.

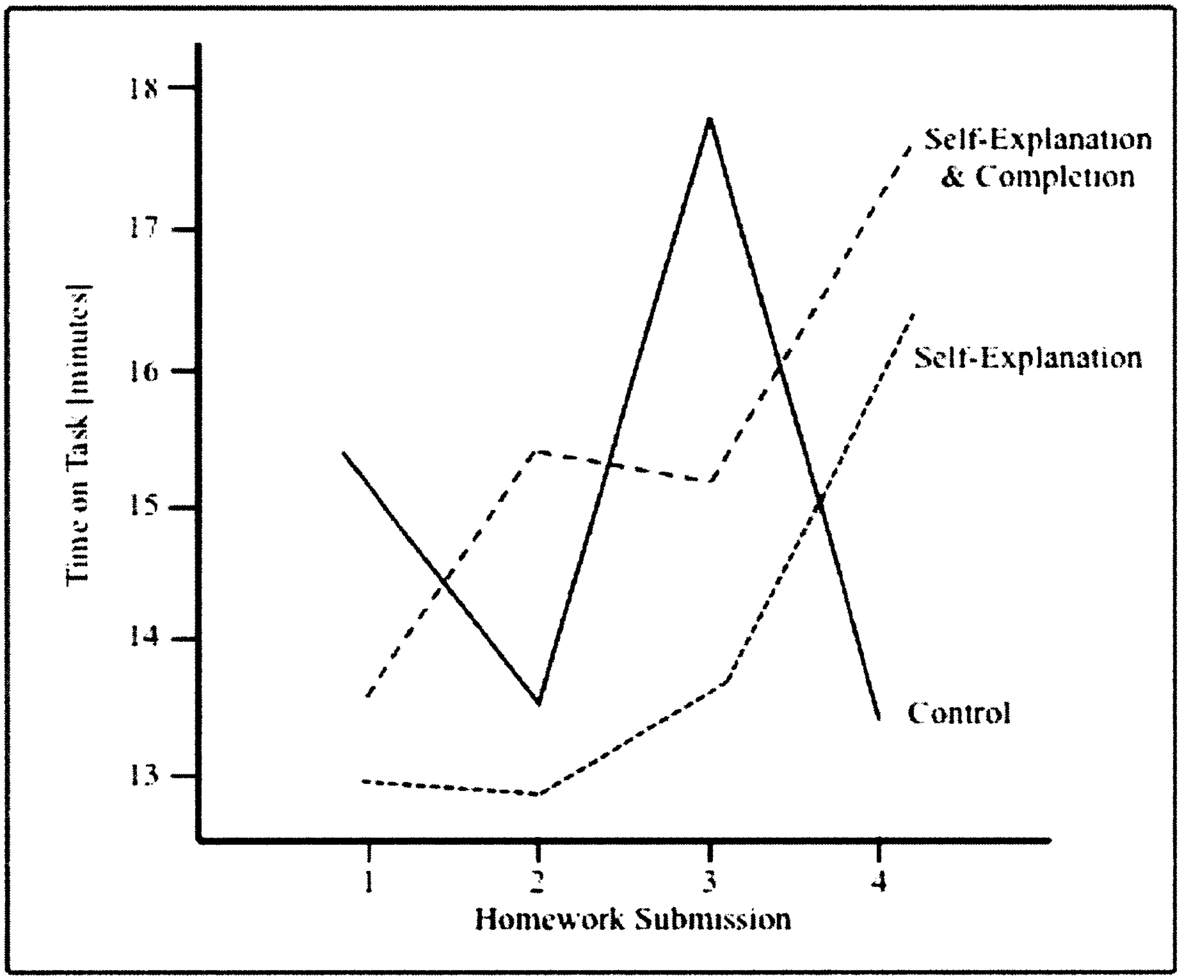


Figure 3. Time on Task for Three Treatment Groups on Four Homework Submissions. Both experimental groups display similar behavior, as submission number increases; yet for the control group, the sign of the slope of time on task changes on each submission to the opposite. However, mean values for the time on task for all treatment groups do not differ significantly.

The following table presents the calculated instructional condition efficiency measures and relative performance and effort scores for the three treatment groups for the four quizzes and the posttest.

Table 5

Instructional Condition Efficiency

Testing event	Instructional condition	Performance	Effort	Relative condition efficiency
Quiz 1	Control	-0.12	-0.29	0.12
	Self-Explanation	-0.04	0.09	-0.09
	Self-Explanation & Completion	0.16	0.20	-0.03
Quiz 2	Control	-0.08	-0.22	0.09
	Self-Explanation	0.33	0.12	0.15
	Self-Explanation & Completion	-0.25	0.10	-0.24
Quiz 3	Control	-0.09	-0.26	0.12
	Self-Explanation	0.31	0.27	0.03
	Self-Explanation & Completion	-0.22	0.01	-0.15
Quiz 4	Control	-0.21	-0.16	-0.04
	Self-Explanation	0.23	0.22	0.01
	Self-Explanation & Completion	-0.02	-0.06	0.03
Posttest	Control	-0.33	-0.23	-0.07
	Self-Explanation	0.27	0.32	-0.04
	Self-Explanation & Completion	0.06	-0.09	0.11

As can be seen from the table, there were two events for which the highest performance resulted in negative efficiency. On quiz one, self-explanation & completion group had the highest average performance z-scores (0.16) but the resulting efficiency was negative (-0.03). Similarly, on the posttest, self-explanation group had the highest average performance z-scores (0.27) but its efficiency was negative (-0.04). There were three events, where the highest performance was not associated with the highest efficiency. On quiz one, self-explanation & completion group had the highest average of performance z-scores (0.16); however, the control group had the highest efficiency (0.12). On quiz three, the highest average performance z-score was in the self-explanation group (0.31) but the highest efficiency (0.12) was again achieved by the control group. On the posttest, although self-explanation group had the highest average of performance z-scores (0.27); the self-explanation & completion group demonstrated the highest efficiency (0.11).

The posttest deserves a more detailed analysis. Figure 2 presents instructional condition efficiency as a perpendicular distance from the performance=effort line.

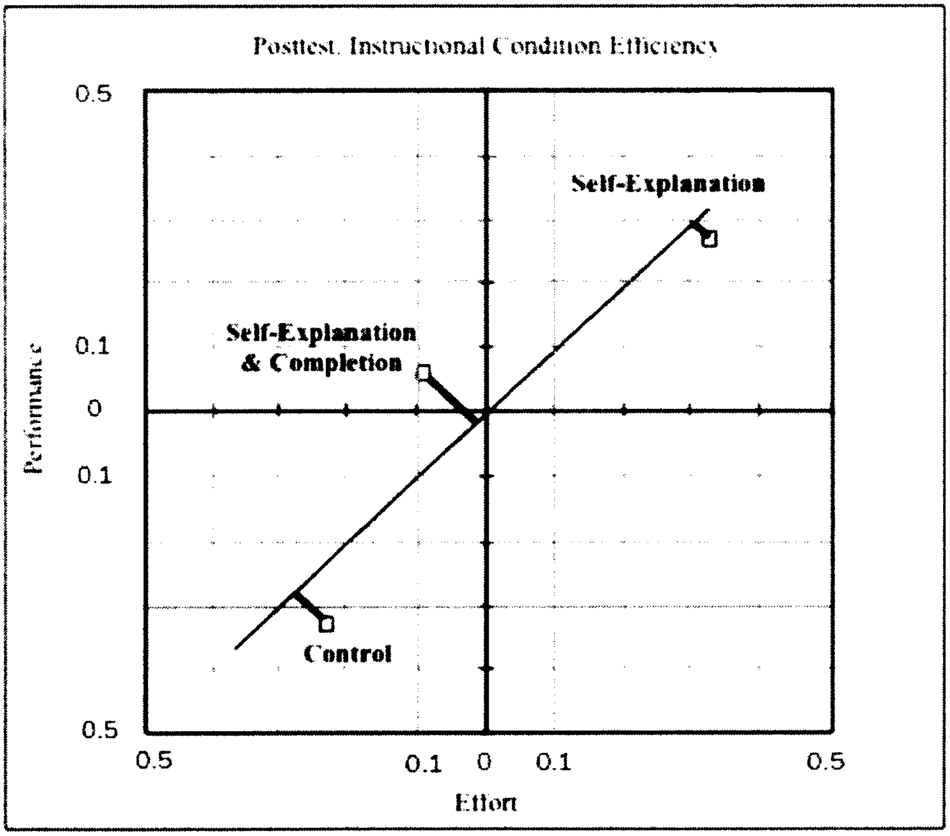


Figure 4. Posttest, Instructional Condition Efficiency.

As graph illustrates, the lowest efficiency is associated with the lowest performance and the lowest effort for the control group. Self-explanation group has the highest performance but also the highest effort, and therefore on the graph is positioned below the line performance = effort (i.e., the efficiency value is negative). Since the posttest performance average z-score of the self-explanation group (0.27) was lower than a posttest effort average z-score (0.32), they resulted in a negative combination performance minus effort, and in turn in a negative efficiency. The instructional condition efficiency formula (as suggested by Paas & van Merriënboer, 1993) implies that better efficiency can be achieved either by the same performance and lower effort, or by the higher performance and the same effort. However, this formula when applied to two extreme situations (i.e., lower left and upper right corners of the graph) leads to unforeseen conclusions. For example in the left lower corner, the control group with the lowest performance and the lowest mental effort produced the highest efficiency (see the first quiz) because the lowest performance was outscored by the lowest effort. On the right upper corner of this graph, self-explanation group produced negative efficiency on the posttest because the highest performance was outscored by the high mental effort. As a result, the self-explanation group was not the most efficient on a posttest as compared to self-explanation & completion group that demonstrated intermediate performance and intermediate effort.

In summary, the results do not support the third hypothesis. Experimental treatments did not decrease time on task and increase instructional condition efficiency understood as better performance in combination with lower mental effort.

Research Question I

The *first research question* asked if the quality of self-explanations changes throughout the duration of the study for the students in two experimental conditions (i.e., prompted self-explain and prompted to self-explain completion problems).

A two-way ANOVA (between subjects-group, within subjects-submission number) on the quality of self-explanations produced a non-significant effect of the group, F(1, 140)=0.49, p=0.48, $\eta^2 = .01$, with the quality score for self-explanation group (M=16.30, SD=4.87) and quality score for self-explanation & completion group (M=15.75, SD=4.90). The effect of interactions was not significant, F(3,140)=0.32, p=0.82. However, the effect of submission was significant, F(3,140)=3.10, $p<.05$, $\eta^2 = .01$. Since homogeneity of variances was assumed, Levene's test F(7,140)=4.86, $p<.001$, a post-hoc Turkey HSD test was used to compare qualities of all participants across submissions and revealed that the first submission (M=14.32, SD=5.97) had a significantly lower score than the last submission (M=17.54, SD=3.67). To further investigate whether the quality of self-explanations improved differently over time, an ANCOVA was used to compare quality of self-explanations scores between two experimental groups across four homework submissions. The independent variable was a quality score, fixed factor was group (two levels), and submission number (four levels) was treated as covariate. Self-explanation & completion group served as a baseline for comparison for the parameter estimates. Levene's test did not allow to reject the null hypothesis, F(1,146)=0.49, p=.53. The effect of submission was significant, F(3,144)=9.45, $p<.01$, $\eta^2 = .06$; the effect of group was not significant, F(1, 144)=1.35, p=.25, $\eta^2 = .01$; and the effect of interaction group x submission was not significant, F(7, 144)=0.92, p=.34, $\eta^2 = .01$. The parameter estimate for submission, B=1.41, t=2.82, p<.01, shows that the overall quality of-self-explanations significantly improved with time. The parameter estimate for group by submission interaction, B= -0.67, t=-0.96, p=.34 reveals that the quality of self-explanations in self-explanation group changed more slowly with time than the quality of self-explanations in self-explanation & completion group, but this effect was not significant. In particular, for self-explanation group, rate of change of the quality of self-explanations with time was 0.74, in contrast to 1.40 (rate of change for self-explanation & completion group). Figure 3 illustrates how quality of self-explanations changed with time for the two experimental groups.

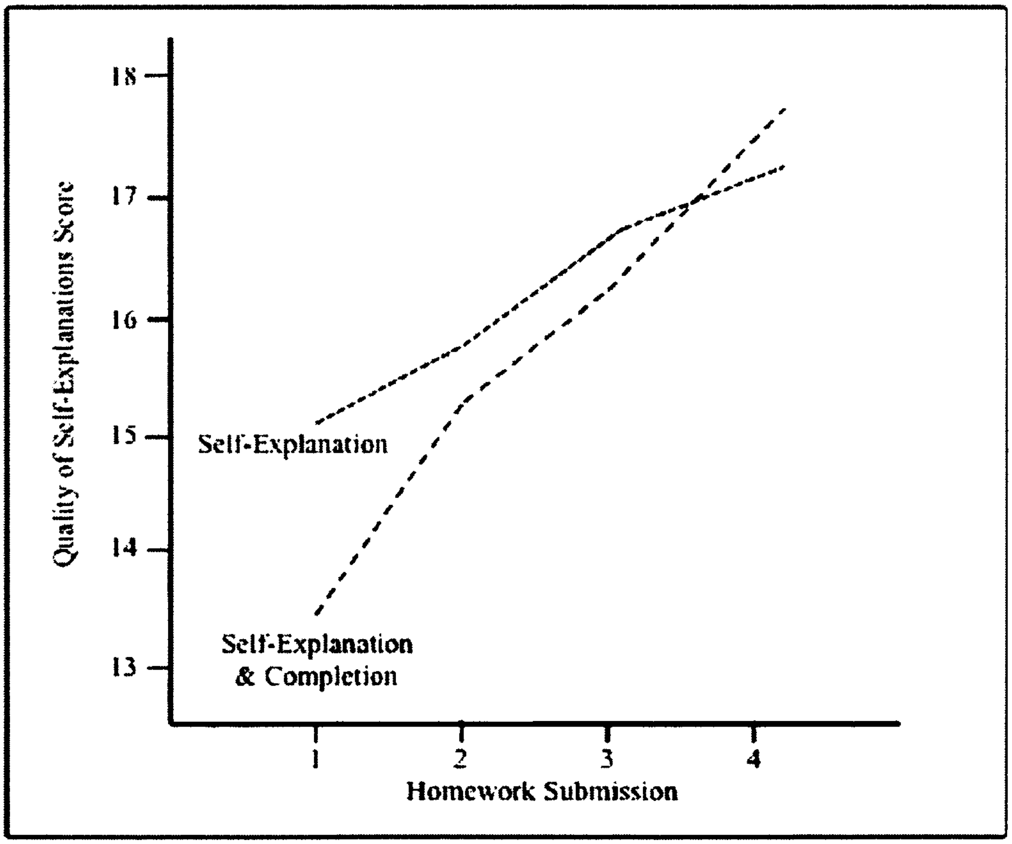


Figure 5. Quality of Self-Explanations for Two Experimental Groups on Four Homework Submissions

Finally, a robust Welch F-ratio test utilizing one-way ANOVA was used to understand whether each experimental group significantly improved quality of self-explanations for the last submission compared to the first. This effect was more pronounced for self-explanation & completion group, Welch F(1, 24.62)=6.58, $p<0.05$, than for self-explanation group, Welch F(1, 33.45)=1.85, $p<.05$, with both effects being statistically significant.

In summary, the quality of self-explanation score was not significantly different between the two experimental groups but the effect of submission was significantly different with both groups significantly improving such quality with time.

Research Question II

The *second research question* asked which treatment group would show a more favorable attitude towards the instructional strategy was answered using an attitude survey. A total of 34 participants completed the survey. Out of those participants, 9 were assigned to the control group, 12 were assigned to prompted self-explain group, and 13 were assigned to prompted self-explain completion group. Survey responses to eight items were analyzed quantitatively. A one-way ANOVA returned a significant effect of group, F(2, 21)=8.00, $p<.01$, $\eta^2 = .34$. Since homogeneity of variances was violated (Levene F(2,21)=2.83, p=0.08), a Bonferroni post-hoc comparison test was used and revealed that control group (M=3.58, SD=1.35) responded significantly different (i.e., was neutral) than self-explanation group (M=2.01, SD=0.45) and self-explanation & completion group (M=1.98, SD=0.516) that on average responded in favor of the instructional strategy (i.e., agreed to eight survey items). No significant difference was found between experimental group responses. The following table presents the summary of participants' responses to these survey items using a 7-point Likert type scale from "Strongly agree" to "Strongly disagree."

Table 6

Summary of Survey Responses

	Average Response		
Survey item	Control (n=9)	Self-Explanation (n=12)	Self-Explanation & Completion (n=13)
I did homework differently compared to how I did it before the study	6.1 (Disagree)	2.8* (Slightly agree)	2.2* (Agree)
The way I approached my homework made me think in-depth about problems	2.2 (Agree)	1.6 (Agree)	1.8 (Agree)
Using the suggested approach helped me to prepare for quizzes and tests	3.5 (Neither agree, nor disagree)	1.8* (Disagree)	1.6* (Disagree)
As I gained experience, I implemented the approach during testing	4.1 (Neither agree, nor disagree)	1.9 (Agree)	2.6 (Slightly agree)
I will continue with the approach after the study is over	3.7 (Neither agree, nor disagree)	2.3 (Agree)	2.7 (Slightly agree)
The approach should be used for teaching physics	4.3 (Neither agree, nor disagree)	1.8* (Agree)	2.1* (Agree)
I had positive experience with the introduced approach	2.9 (Slightly agree	2.4 (Agree)	1.4 (Strongly agree
I benefitted from the study	1.8 (Agree)	1.5 (Agree)	1.4 (Strongly agree)

In addition, separate analyses for each survey item were performed utilizing one-way ANOVA. Note, * represents significantly different experimental group responses compared to the control group responses at $p<.05$. No significant differences were found between two experimental groups.

The next survey items solicited participants for comments. When asked, how suggested during the study approach to solving problems helped them learn, participants in three treatment groups responded differently. In the control group, five participants explained that it was helpful to write down their solutions on paper, and one participant complained about being assigned to control group without any hints or assistance. In self-explanation group, one participant complained about being forced to do more work than usual, the rest of the participants indicated that the approach helped them in the following ways: “checking units helped,” “explain how to crack problems,” “forced to think more focused,” “allowed to understand the theory beyond the given events,” “made me understand concepts,” “showed a different thought process,” “helped to do well on tests.” In self-explanation & completion group, one student responded pessimistically, “Nothing helps me!!!” Five others indicated partial solutions as the most helpful part of the instructional approach. The following are examples of the comments related to self-explanations in this group: “helped me to break down steps in workable parts,” “being forced to explain made me focus more,” “slow down and write down all the equations made me take time to think.”

When asked to comment on implementing the suggested instructional technique during testing, in the control group, only two people commented. They explained that they implemented the suggested approach because writing down solutions helped them to organize their thoughts and plan. In self-explanation group, seven participants provided comments. The following are examples of their comments: “made me think critically,” “made me think longer and deeper,” “I began with thinking about equations for each topic,” “solutions came naturally to me.” In self-explanation & completion group, also seven participants commented. The following are examples of their comments: “I used knowledge from partial solutions,” “I focused more on formulas then on numbers,” "having formulas before starting made it easier to understand," "I took time to read problems and think about them critically."

When asked to comment on implementing the suggested approach to solving problems in the future, in the control group, there were three comments where participants indicated that they were assigned to the control group, and therefore they were not exposed to any specific approach. However, six participants plan to continue implementing the suggested approach which they understood as writing down their solutions on paper. They explained that it helps their thought process. In self-explanation group, there were unexpected comments stating that recording time on task and mental effort should not be carried over. In this group, five participants do not plan to continue using the suggested approach because they consider it time consuming, although three of those who do not plan indicated that they would do it in their heads. Seven participants in this group plan to continue using the approach. They explained that it helped them improve quiz- and test- performance and increased understanding of the subject. In addition, two people indicated that they plan to use it in other classes. In self-explanation & completion group, there were four comments that can be summarized similar to: "I would like to implement but where would I get partial solutions?" In this group, three people do not plan to use the approach because it requires a lot of effort and is time consuming, the rest plans. From those who plan, comments divided in favor of partial solutions, similar to: "I wish I could continue but I need partial solutions," and in favor of self-explanation technique, characterizing this approach as: "organized way of learning," "helps to break problems down into manageable chunks", "allowed better understand what I was doing," "made me think critically."

When commenting to the question of whether the approach should be used for teaching physics, the control group mostly put N/A with one participant commenting in favor of the approach in the following way, "teachers should know how much time and effort it takes students to do work at home, they should reevaluate teaching methods!" All participants in self-explanation group responded that the suggested approach should be used for teaching physics providing the following comments: "efficient way of study," "unit check helps a lot," "shows students what they do not understand," "focus on steps not on numbers helps in other classes too," "people aren't that fast mentally, helps to see details that matter," "helps to organize thoughts as it helped me." In self-explanation & completion group, one student commented that approach should not be used for teaching physics because it helps only those who do poor but takes too much time from those who do well anyway; and those who responded in favor of the approach provided the following comments: "helps to work problem into parts-makes a lot of sense to do it this way," "helps to recognize and connect all the elements," "makes it easier to understand the problem," "helps to start even if you don't know how to solve it in the beginning."

With regard to the satisfaction with the study and the suggested instructional technique, the following are representative comments to the survey item that asked to name one aspect participants especially liked about the introduced approach to solving problems. In the control group: "having to thoroughly write out the problem," "having to write out the thought process on paper allowed me to follow over my work again later." In self-explanation group: "slow down and understand the problem before putting numbers," "to look back and correct mistakes," "write down equations before plugging in numbers," "allowed me to start the work, the rest came naturally," "explaining to myself how I get the answers allowed me to think what I was doing instead of doing without any approach," "checking the units after solutions-wonderful technique!" In self-explanation & completion group, five people mentioned partial solutions similar to the comment, "having solutions before solving helped a lot." The following are examples of the comments from this group that were not related to partial solutions: "I was pushed to make sure that I knew what I was doing," "helped to be more organized," "think critically," "monitor myself," "breaking down solutions into steps made problems less daunting."

When participants were asked to name one aspect about the suggested approach to solving problems or the study, participants in all three groups mentioned recording time and effort; in the control group it was mentioned twice; in self-explanation group, eight times; and in self-explanation & completion group, five times. No other negative aspects were mentioned except one student in self-explanation group has commented that being bossed by someone else at home was annoying.

Last, all participants positively reacted to the study with comments appreciating an opportunity to be introduced to instructional technique with three comments asking to continue with providing partial solutions and five comments asking to take away recording mental effort and time on task.

Summary

In summary, all treatment groups positively reacted to the study and to the instructional technique they were introduced to. Participants in the control group mostly understood this instructional technique as writing down solutions on paper that has helped to organize their thought process. Self-explanation group pointed out the advantages of self-explanations technique as forcing learners to think in-depth about problems and guiding on how to break down solutions into manageable steps. Within self-explanation & completion group, participants either saw the advantages of the presented to them instructional technique in having partial solutions that allowed them to channel their thoughts into the right direction, or being prompted to self-explain with the same advantages that were pointed out by participants in self-explanation group.

CHAPTER IV

DISCUSSION AND CONCLUSIONS

Several findings of this study were contrary to the predictions. On one hand, results of this study suggested limitations of cognitive load theory' on the other hand, they suggested a different understanding of the relationship between performance and total cognitive load when the goal of instructions is teaching complex tasks. The discussion focuses on quantitative results; while qualitative results help interpret the findings. The discussion starts with a closer look at performance because it is the main goal of instructions.

Hypothesis 1I: Participants in two experimental conditions will perform better than participants in the control condition on quizzes and a posttest.

For the quizzes, the effect of treatment was near significant for the self-explanation group as compared to the control, with self-explanation & completion group not being significantly different from self-explanation group or the control. For near transfer quiz performance groups did not perform differently. As for far transfer quiz performance, self-explanation group significantly outperformed both the control and self-explanation & completion group. The self-explanation group demonstrated a better performance on the posttest as compared to the control; however self-explanation & completion group outperformed the control only on near transfer posttest questions. Next three figures summarize results of this study related to performance.

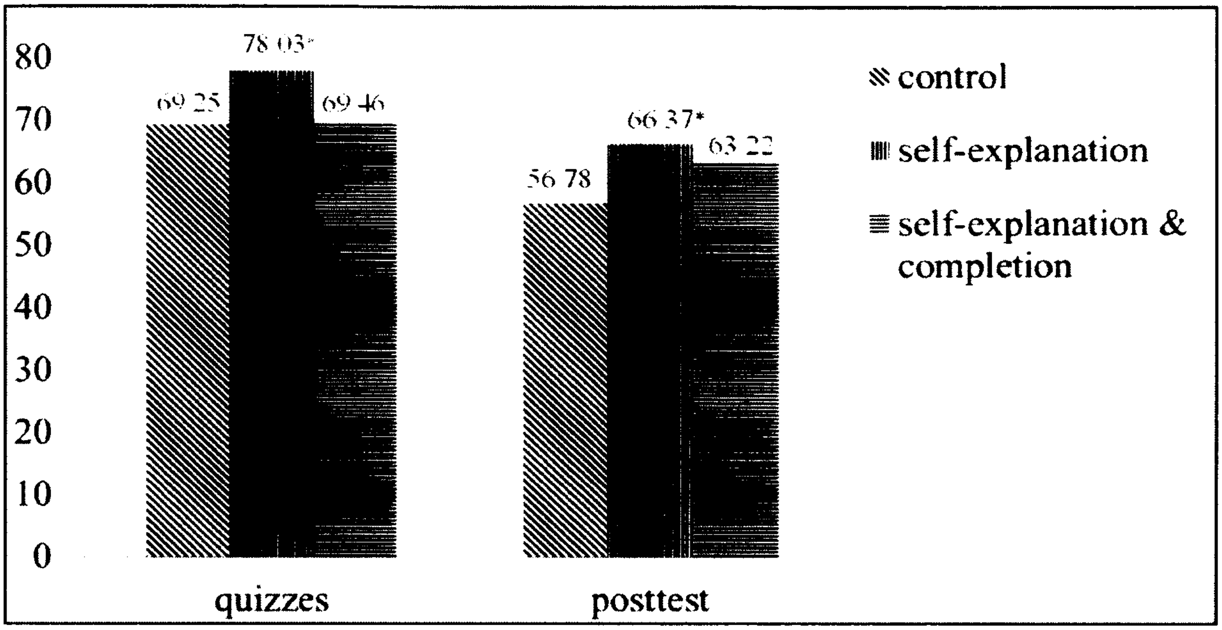


Figure 6. Quiz and Posttest Grades. Note. * represents significantly different results as compared to the control group at p<.05.

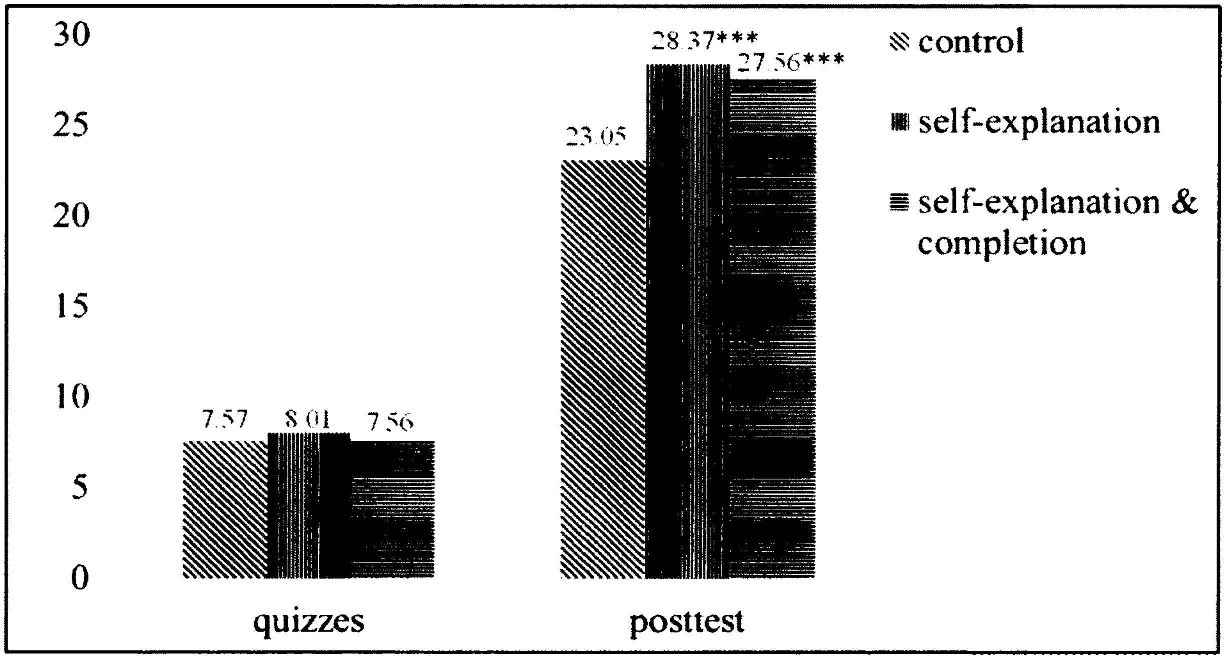


Figure 7. Near Transfer Scores. Note. *** represents significantly different results as compared to the control group at p<.001.

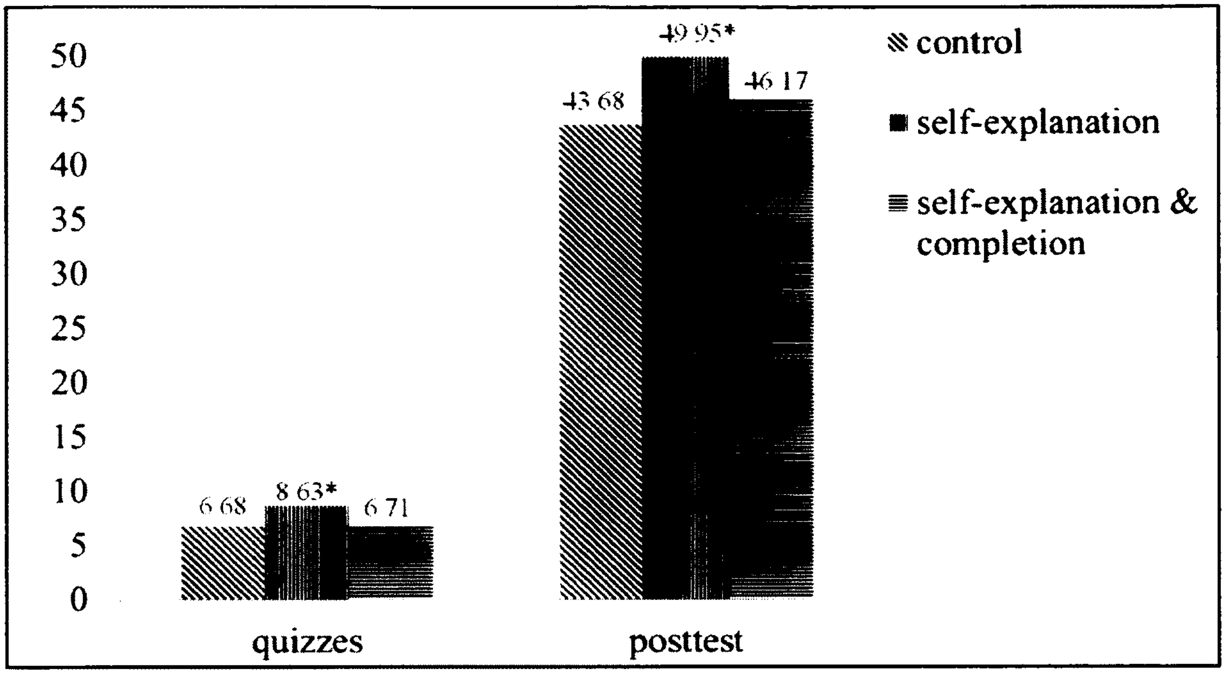


Figure 8. Far Transfer Scores. Note. * represents significantly different results as compared to the control group at p<.05.

The combination of self-explanations and completion problems failed to deliver a better performance especially on far transfer needs further attention. As far transfer requires generalization, it allows learners to apply skills in contexts that are different from those encountered during instructions. Similar results (i.e., significant effects of treatment for far transfer test performance, but not for near and intermediate transfer test performance) were found in Van Merrienboer's et al. (2002) study where researches investigated training of complex cognitive skills and effects of manipulations with cognitive load by redirecting attention from extraneous to germane processes on transfer of performance. These results may be explained by the assumption that instructions based on CLT typically have the strongest effect on far transfer because better organized cognitive schemas are primarily useful when learners must deal with new situations (Paas & van Merrienboer, 1994; van Merrienboer, et al., 2002). Self-explanations could be especially important for solving far transfer problems because as learners explain to themselves the rationale for the solution steps they better understand how to apply domain principles and achieve goals using certain operators (Chi et al., 1989). In this study, students in self-explanation & completion group were prompted to self-explain. Unlike self-explanation group, self-explain and completion group had partial solutions in addition to the self-explanations. This strategy combination suggests that partial solutions might have been distracting from the self-explanation activity. During the learning phase, participants in self-explanation & completion group saw partial solutions, they immediately directed their attention to completing those solutions (i.e., focused on understanding how to derive answers to the problems in algebraic form) rather understanding the solution through the use of self-explanations. In contrast, in self-explanation group learners started their thought process with identifying the topic and the related general principle. This use of both strategies may have confused the self-explanation & completion group during the testing phase where learners did not have help from partial solutions and had to either recall how they solved the familiar problem (near transfer) during the learning phase or start completely from scratch in unfamiliar situation (far transfer). Practicing near transfer problems on the homework was helpful for the self-explanation & completion group and resulted in better test performance as compared to control group. For the far transfer, during the testing phase learners the in self-explanation & completion group had to start solving the problem by identifying the topic related to this problem and recording the formula for the general principle for this topic. This task was not practiced at home in the same manner because identifying the topic and the general principle was always easier during the learning phase since partial solutions presented this information. For example, students had to solve the following homework problem, "Blood with density

1.06 g/cm^3 and 10-kPa gauge pressure flows through an artery at 30 cm/s. It encounters a plaque deposit where the pressure drops by 5 %. What fraction of the artery's area is obstructed?" When learners first read this problem, they could be not sure what topic it relates to and what major principal needs to be applied to solve it. However Bernoulli's equation, i.e., $p + \frac{1}{2}\rho v^2 = p' + \frac{1}{2}\rho v'^2$, presented them a partial solution and likely enabled them to recall that this equation relates to fluids. Next, they could consult their textbook or class notes and report that they are investigating fluid motion, in particular fluid dynamics, and the problem requires application of conservation of fluid energy in the form of Bernoulli's equation that reads that total energy per unit volume of fluid is conserved as the fluid moves.

Identifying the topic related to the problem and the general principle is the key aspect of solving any physics problem, which becomes especially important during comprehensive test where problems represent multiple topics. Survey responses where self-explanation & completion group had a split opinion in identifying the main idea of the instructions may suggest support for this interpretation. Within this group, some participants reported the benefits of the instructional approach used to solve homework problems that included partial solutions. Other participants in this group reported that they mostly benefited from being prompted to self-explain.

Hypothesis I: Perceived cognitive load imposed on participants in the two experimental conditions (i.e., prompted to self-explain and prompted to self-explain and complete) will be lower than perceived cognitive load imposed on participants in the control condition during both learning and testing phases.

The following figure summarizes results for cognitive load measures during homework, quizzes, and a posttest.

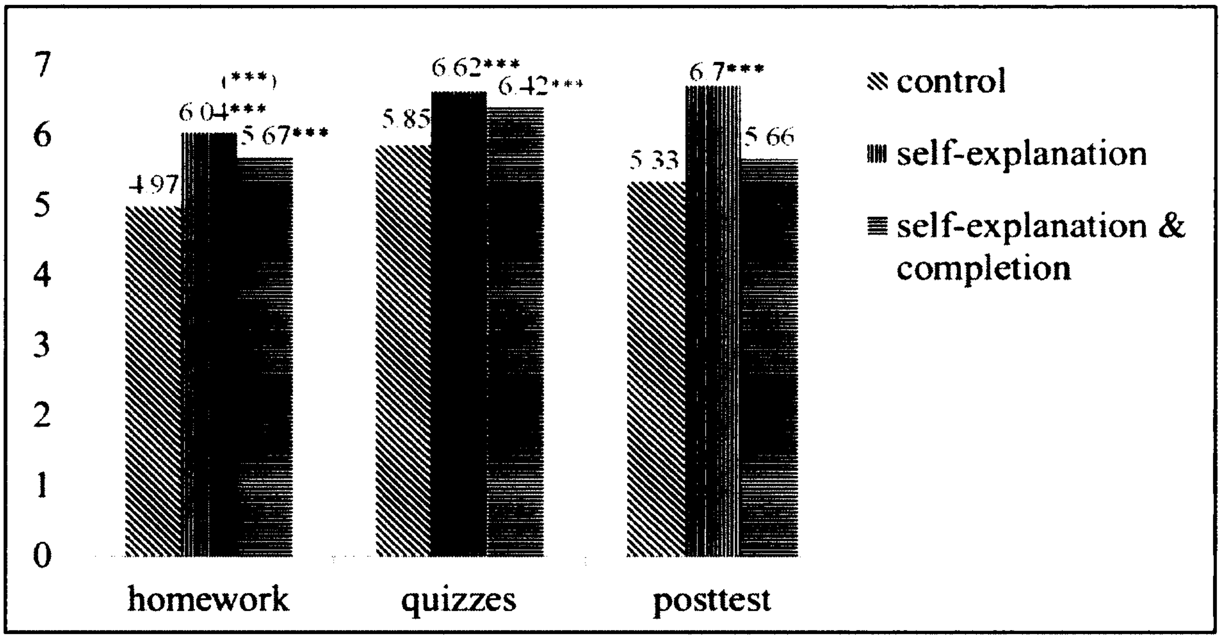


Figure 9. Cognitive Load Measures. Note. *** represents significantly different results as compared to the control group at p<.001.

The second unexpected result is that higher cognitive load was associated with better performance. For this study, all three treatment groups were initially presented with instructions that reduced extraneous and intrinsic cognitive load. Since learning for all treatment groups occurred, it can be assumed that total cognitive load stayed within limits of working memory capacity. During the instruction, participants' working memory had enough capacity that could be redirected and used as germane resources. During the learning phase, for the control group, the utilization of germane resources was low and resulted in the lowest reported total cognitive load. In self-explanation & completion group, the total reported load was higher than reported by the control group because germane processing was induced by self-explanation activity and extraneous load was further decreased by the presence of completion problems. In the self-explanation group, the reported total load was the highest, and could be attributed mainly to germane processing. For the testing phase, analysis of total cognitive load presented by van Goh and Paas (2008) was used as a basis for the prediction in this study. They contrasted the learning and testing phases expecting lower cognitive load reported for the test as compared to the learning phase for the experimental conditions that aim to simultaneously reduce extraneous cognitive load and increase germane processing. However, the results of this study indicated that contrary to prediction during the testing phase, complex problems were rated higher in invested mental effort for the two experimental groups as compared to control. This situation can be illustrated with one story-problem posttest question. The problem was stated in the following way: "Jill has gotten out of her car in the Wal-Mart parking lot. The parking lot is on the hill and has a 5° slope. Twelve meters downhill from Jill, a tiny old lady lets go of a fully loaded 10-kg shopping cart. The cart, with its frictionless wheels, starts to roll straight downhill. Jill immediately starts to sprint after the cart with a top acceleration of 1.8 m/s^2. a) What is the acceleration of the cart? b) How much time has elapsed before Jill catches the cart? How far has the cart rolled?" In the control group, five participants rated this question for the mental effort as low as 3 but got partial credit of 2 points out of 10 possible. The rest rated the mental effort as high as 7, 8, or 9. , with performance varying from 0 to full credit of 10 points. In self-explanation group, only two participants rated it as intermediate effort of 5 and both received maximum points possible for performance; the rest rated it as high as 8 or 9, with performance varying from 3 points to full credit. Similarly, in self-explanation & completion group, three participants rated this question as low-to-intermediate effort of 4 and received the maximum points possible with most other group member ratings in a high category of 7 or 8 one participant rating it as high as 9 while performance differed from 2 points to full credit. The following table shows average mental effort ratings, used in this study as measures of total cognitive load, and average performance (i.e., awarded out of 10 maximum points, see Appendix E) for the three treatment conditions.

Table 7

Example of Story-Problem Mental Effort Rating and Performance by Treatment Group

Group	Mental Effort	Performance
Control	6.29	6.11
Self-Explanation	7.67	7.79
Self-Explanation & Completion	7.08	6.73

This finding suggests that participants who rated the problem as requiring low mental effort in the control group could be contrasted with participants in the two experimental groups who rated this problem as requiring low-to-intermediate mental effort because low ratings in the control group were combined with poor performance but low-to intermediate ratings in the experimental groups were combined with maximum points possible. Based on this example, it is reasonable to assume that both experimental groups used their germane resources that may have been prompted by self-explanations more so than those in the control group. The experimental treatment participants were better able to understand the question and what it takes to solve the problem, and therefore solve it correctly. In contrast, the control group was not able to adequately assess what was needed to solve the problem, and consequentially underestimated mental effort and overestimated quality of their solutions. The higher mental effort in this case for both experimental groups can be attributed to better utilization of germane cognitive resources. The self-explanation group was more successful than self-explanation & completion group in terms of solving the problem correctly because self-explanation group was acting in the familiar situation, since they practiced at home under the same conditions. The self-explanation & completion group was put in an unfamiliar situation possibly because their homework had provided partial solutions. But, overall both experimental groups were better able to understand the problem and utilize their germane resources as compared to the control.

Hypothesis III: The two experimental conditions will have higher efficiency than the control condition.

The third unexpected result of this study was that comparison of instructional condition efficiency for the three treatment groups produced inconclusive results. These findings raise a concern about the application of the formula that is widely used for efficiency. de Jong (2009) reported that low performance associated with low cognitive load would result in a similar efficiency value as that of high performance with a high cognitive load, which leads to unclear experimental situations. Similarly, Moreno and Valdez (2005) reported equivalent efficiency for two groups where one group was presented with a set of frames in the right order and the other group that had to put those frames in order themselves. Moreno and Valdez concluded that the experimental testing of efficiency produced unclear results when one group with higher performance and higher cognitive load combination was not different from a group with lower performance and lower cognitive load combination.

Because efficiency formula, as a combination of performance and cognitive load, returned inconclusive results, it may be more beneficial to use Morrison's et al. (1988) or Tennyson's (1980) approach for the future research. In these studies, learning efficiency was understood as a combination of performance and time on task. However, such an approach would be much easier to implement in CBI because the system would be capable of recording time on task automatically.

The model suggested by Paas and van Merrienboer (1993) uses graphical representation of efficiency. It assumes that the experimental treatments are targets of the design based on CLT and should be above the line "performance = effort." However, as mental effort increases, it assumes that performances should increase at least linearly and there is no indication on the line that at some point overload occurs. The line goes up indefinitely, however, in reality it needs to have some point at which learning should slow down with increase in mental effort as an indication of a zone where overload is likely to occur and then a steep down turn where further increase in cognitive load is detrimental to learning, representing cognitive overload. Such curve can be found in the psychology literature, but it describes performance as a function of arousal. The Yerkes-Dodson law (1908) states that increase the arousal level of the individual increases performance up to an optimal level beyond which over-arousal leads to deterioration in performance. The law also states that such deterioration occurs more quickly when the task to be performed is complex. Yerkes-Dodson law can be illustrated with the following graph.

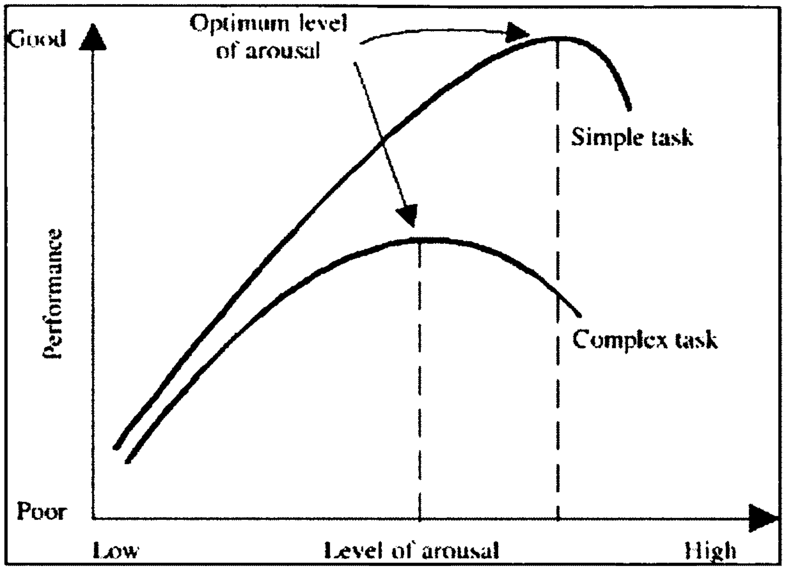


Figure 10. Illustration of Yerkes-Dodson Law (1908). Adapted from "Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents: Part 1: Overview of the IDAC Model," by Y .H .J. Chang and A. Mosleh, 2007, *Reliability Engineering and System Safety, 92*(8), 997-1013.

Arousal is defined as a change in physiological and/or psychological responsiveness to internal or external stimuli (Howells, Stein, & Russell, 2010). Howells et al. (2010) further explain that difficulty, complexity and stress-inducing tasks lead to increased subjective perceptions of mental effort which in turn can been related to increased physiological arousal, and therefore subjective perception of mental effort may reflect changes in arousal during performance of attentional tasks. Based on this explanation, self-reported mental effort measures for this study can be considered a reflection of changes in arousal during performance of a cognitively demanding complex task. The next graph illustrates the curve resulting in connecting efficiency measures for the three treatment condition for the study.

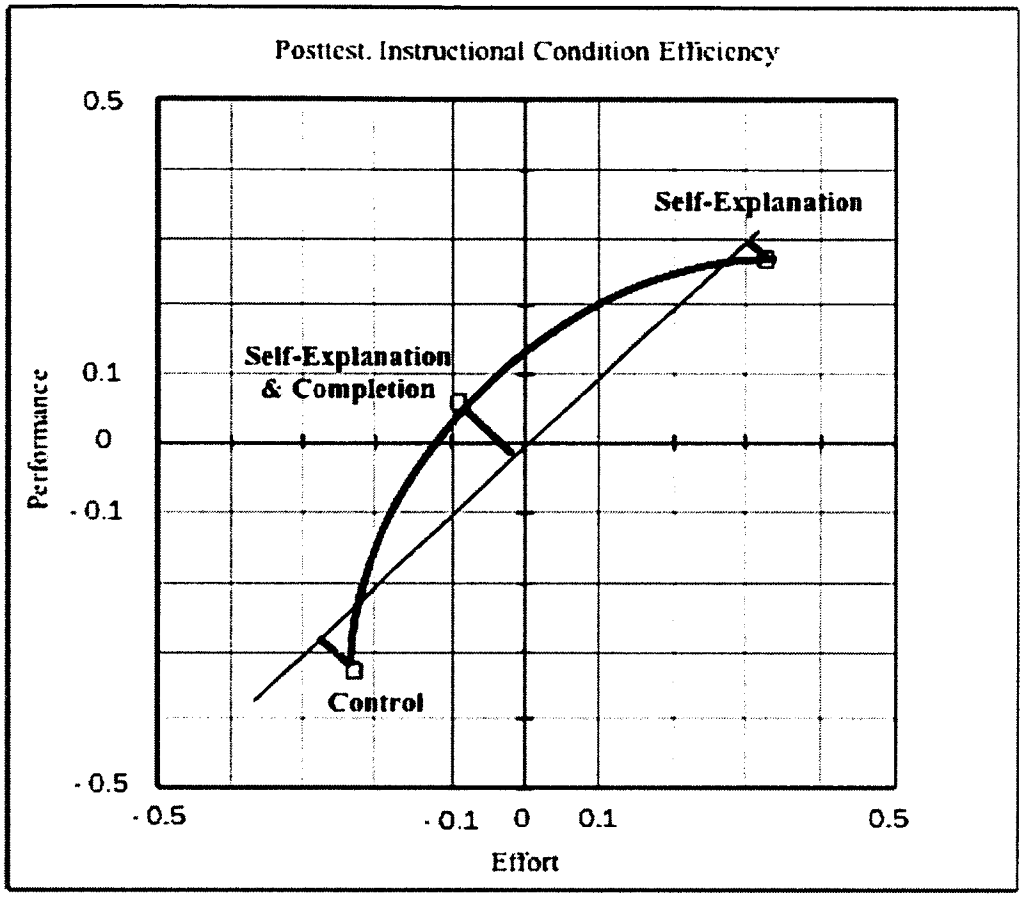


Figure 11. Curve Connecting Efficiency Measures for Three Treatment Groups.

Since graph illustrates relative efficiency, the control group can serve as a baseline. Moving the curve, so that the control group has the lowest efficiency as a combination of the lowest performance and lowest mental effort, in the first quadrant of the Cartesian coordinate system; the resulting graph illustrates the behavior similar to the Yerkes-Dodson law.

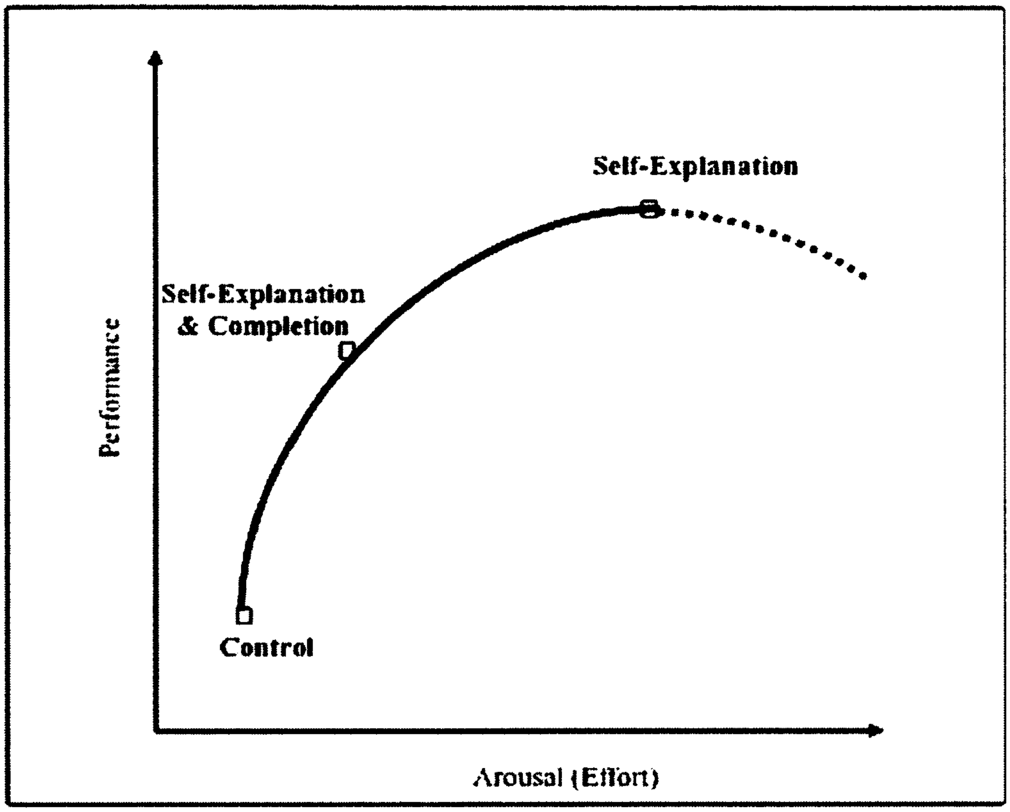


Figure 12. Curve Connecting Efficiency Measures for Three Treatment Groups.

As seen in this graph when mental effort increase, it leads to a better performance, but only to a certain point on the graph where an additional increase in mental effort slows down the increase in performance. At the point where mental effort is the highest (self-explanation group), the performance is also the highest; however, further increases in mental effort may lead to cognitive overload and may prevent further learning.

The suggested graphical representation could serve as a better model for understanding the relationship between mental effort and performance because it takes in consideration non-linear relationship between these two measures and has a reflection on the graph for cognitive overload and a situation that precedes this overload. The model behind Yerkes-Dodson law allows us to explain the relationship between total cognitive load and performance. Based on this relationship it is not detrimental, but beneficial for instruction to increase total cognitive load to the point that precedes overload, but keep it within limitations of the working memory. This explanation leads to a clearer formulation of the major consideration for the design of instructions when teaching complex tasks from a CLT perspective: in the initial stages of instructional presentations, minimize extraneous and intrinsic load to allow working memory allocate resources for the germane processing; and on the later stages, impose techniques that induce utilization of germane resources, so that this utilization is maximized for the specific task but total load stays within limits of the working memory.

Efficiency can also be understood as reducing time on task. The lack of significant differences for the time that learners invested in solving problems in this study contradicts learners' perception of how much time they have spent. In their survey responses, participants in two experimental groups frequently mentioned that they saw the downside of the suggested instructional technique as requiring more time. However, results indicate that such time on average was not different between groups.

Research Question I: Will the quality of self-explanations change throughout the duration of the study for the students in two experimental conditions (i.e., prompted self-explain and prompted to self-explain completion problems)?

The last, analyses of the quality of self-explanations suggest that the quality of the explanations may improve with time. However, the quality of self-explanations that significantly improved in both experimental groups with time deserves a more detailed look. Participants in this study major in the following fields: computer science, chemistry, biology, mathematics, physics, and engineering. Prior research (Chi et al., 1989; Renkl, 1997) suggested that learners differ considerably in their ability to self-explain. For this study, it was decided to divide learners into three groups: low, with the quality of self-explanations ranging [1-14]; intermediate, ranging [15-19]; and high, ranging [20-21]. Table 7 shows the number of participants in each sub-group on the first and the last submission, and the change in the number of participants in each sub-group for the two experimental groups.

Table 8

Changes in Number of Participants within Quality of Self-Explanations Sub-Groups

	Self-Explanation Group			Self-Explanation & Completion Group		
Submission/Quality	low	intermediate	high	low	intermediate	high
1st Submission	8	3	8	8	3	7
4th Submission	2	7	10	2	8	8
Change	-6	+4	+2	-6	+5	+1

The table illustrates changes between sub-groups were similar for the two experimental treatments. It also can be seen that major changes occurred between low quality and intermediate quality sub-groups; the minimal changes occurred in the high quality sub-group. It appears that differences between participants' ability to self-explain had a great impact on the changes, and participants primarily improved their quality from low to intermediate but not from intermediate to high quality. Similar information could be retrieved from the results for the main effect of the group based on two-way ANOVA, F(1, 140)=0.49, p=0.48, $\eta^2 = .01$. Partial eta squared of 0.01 indicates that the model is weak and only 1% of variance of the quality of self-explanation scores can be accounted for group. This finding could be an indication of a high impact of each individual's ability to self-explain. One possible interpretation of the results of this study could be that although the quality of self-explanations for both experimental groups improved significantly with time, individual participants varied significantly in their ability to self-explain. And a high quality of self-explanations may be an individual characteristic of the learner.

The other possible interpretation of the results of this study could be based on prior research (Johnsey, Morrison & Ross, 1992; Weinstein, 1982) suggesting that it may take longer for students to learn how to use generative strategies. In this study, learners were exposed to instructional techniques for a period of 6 weeks. For the future research, it may be beneficial to conduct studies for even longer periods of time or have extensive training on how to do self-explanations prior to treatments in order to understand whether learners who were providing low quality self-explanations at the beginning of the study can improve their quality self-explanations to high quality. Since specific feedback on self-explanations was not provided in the current study, perhaps providing feedback on self-explanations might improve self-explanation quality over time.

Duration of this study over a period of six weeks deserves special attention. As Clark and Snow (1975) pointed out, the amount of time students spend in treatment may significantly affect the outcomes of studies when research is concerned with methods of designing instructions with the goal of presenting them in a classroom. Long-term interventions with learners' exposure to treatments over several weeks can be critical in order to generalize the outcomes of the study and to draw conclusions about the effectiveness of the suggested design of instructions. Clark (1985) argued that a "novelty" effect due to learners' initial enthusiasm to learn more with the newly introduced courseware may produce significant results when the study lasts four weeks or less with this effect dissipating when study lasts between five and eight weeks, and becoming even less pronounced when the study duration is eight weeks or more. This study produced strong effects with regard to cognitive load and performance and lasted six weeks. Prior research (Johnsey, Morrison, & Ross, 1992; Weinstein, 1982) suggests that lengthy periods of time are needed to teach learners how to use generative strategies. Therefore, it may be suggested that for the current study's duration allowed, on one hand, to give learners substantial time to learn how to use self-explanation strategy. This finding is supported the by significantly higher quality of self-explanations at the end of the study compared to the beginning of the study for both experimental treatments. On the other hand, significant difference between treatments obtained for a six week intervention suggests that the produced effect cannot be attributed to the novelty effect.

Conclusions

In conclusion, the suggested instructional technique (i.e., prompting learners to self-explain while independently solving problems) was tested in a long-term study in an authentic environment and produced better performance than conventional problem-solving including both near and far transfer questions. A combination of self-explanations and completion problems demonstrated a better performance only on near transfer posttest questions as compared to control. For all testing events, better performance was associated with higher measures of total cognitive load. It was also determined that self-explanations with and without completion problems do not require additional time on task, and learners may improve quality of self-explanations with time. Since some of the results contradict a major premise of CLT that the design of instructions should aim the reduction of total cognitive load, and also that efficient from CLT perspective instructional conditions should attain better performance in combination with lesser mental effort, a different model based on Yerkes-Dodson law was suggested for explanation of the relationship between mental effort and performance. This model allows a more clear formulation for a major consideration for the design of instructions.

The suggested instructional technique based on prompting learners to self-explain while independently solving problems not only has demonstrated an improved performance but has several advantages, such as: (a) it is as easy to implement in a real classroom, (b) it does not require any additional cost, (c) it does not require additional time on task, (d) it allows learners to adopt a systematic approach to problem-solving, (e) it can be used in the other complex domains of knowledge where problem-solving is required, and (e) learners give it positive feedback.

References

Aleven, V., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science*, *26*(2), 147-179.

Atkinson, R. K., Renkl, A., & Merril, M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading. *Journal of Educational Psychology*, *95*(4), 774-783.

Ayres, P. (2006). Using subjective measures to detect variations of intrinsic load within problems. *Learning and Instruction, 16*(5), 389–400.

Beckmann, J. F. (2010). Taming a beast of burden-On some issues with the conceptualization and operationalization of cognitive load. *Learning and Instruction, 20*(3), 250-264.

Clark R. E. & Snow R. E. (1975). Alternative designs for instructional technology research. *AV Communication Review, 23*(4), 373-394.

Clark, R.E. (1985). Confounding in educational computing research. *Journal of Educational Computing Research, 1*(2), 445-460.

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

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*(4), 293–332.

Chang, Y .H .J., and Mosleh, A. (2007). Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents: Part 1: Overview of the IDAC Model. *Reliability Engineering and System Safety, 92*(8), 997-1013.

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. *Cognitive Science*, *13*(2), 145-182.

Chi, M. T. H., DeLeeuw, N., Chiu, M.-H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*(3), 439-477.

Cooper, G., Tindall-Ford, S., Chandler, P., & Sweller, J. (2001). Learning by imagining. *Journal of Experimental Psychology: Applied*, *7*(1), 68–82.

Davis, E. (2003). Prompting middle school science students for productive reflection: Generic and directed prompts. *Journal of the Learning Sciences*, *12*(1), 91-142.

de Jong, T. (2010). Cognitive load theory, educational research, and instructional design: Some food for thought. *Instructional Science*, *38(2)*, 105-134.

DeLeeuw, K. E., & Mayer, R. E. (2008). A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of Educational Psychology*, *100*(1), 223-234.

Howells, F. M., Stein, D. J., & Russell, V. A. (1010). Perceived mental effort correlates with changes in tonic arousal during attentional tasks. *Behavioral and Brain Function*, *6*(39), doi:10.1186/1744-9081-6-39

Johnsey, A., Morrison, G. R., & Ross, S. M. (1992).Promoting generative learning in computer-based instruction through the use of elaboration strategies training. *Contemporary Educational Psychology*, *17*(2), 125-135.

Kalyuga, S., Chandler, P., & Sweller, J., (1998). Levels of expertise and instructional design. *Human Factors*, *48*(1), 1-17.

Kalyuga, S., Chandler, P., & Sweller, J., (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology, 13*(4), 351-371.

Kalyuga, S., Chandler, P., & Sweller, J., (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology, 92*(1), 126-136.

Kalyuga, S., Chandler, P., & Sweller, J., (2001). Learner experience and efficiency of instructional guidance. *Educational Psychology, 21*(1), 5-23.

Kester, L., Kirschner, P. A., & van Merriënboer, J. J. G. (2004). Timing of information presentation in learning statistics. *Instructional Science, 32*(3), 233–252.

Koedinger, K. R. & Aleven, V. (2007). Exploring the assistance dilemma in experiments with Cognitive Tutors. *Educational Psychology Review, 19* (3): 239-264.

Miller, G. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81-97.

Moreno, R., & Valdez, A. (2005). Cognitive load and learning effects of having students organize pictures and words in multimedia environments: The role of student interactivity and feedback. *Educational Technology Research and Development, 53*(3), 35–45.

Morrison, G. R., Ross, S. M., & O'Dell, J. K. (1988). Text density level as a design variable in instructional displays. *Educational Communication and Technology, 36*(1), 103-115.

Morrison, G. R., Ross, S. M., Kalman, H. K., & Kemp, J. E. (2011). *Designing effective instruction* (6th ed.). Hoboken, NJ: John Wiley & Sons.

Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology, 87*(2), 319–334.

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive load approach. *Journal of Educational Psychology, 84*(4), 429-434.

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

Paas, F., Tuovinen, J. E., Tabbers, H., & van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist,38*(1), 63–71.

Paas, F., Tuovinen, J. E., van Merriënboer, J. J. G., & Darabi, A. A. (2005). A motivational perspective on the relation between mental effort and performance: Optimizing learner involvement in instruction. *Educational Technology Research and Development, 53*(3), 25–34.

Paas, F., & van Merriënboer, J. J. G. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors, 35*(4), 737–743.

Paas, F., & van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem solving skills: A cognitive-load approach. *Journal of Educational Psychology, 86*(1), 122–133.

Paas, F., van Merriënboer, J. J. G., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills, 79*, 419–430.

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58*, 193–198.

Pollock, E., Chandler P. & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction 12*(1), 61-86.

Renkl, A. (1997). Learning from worked-out examples: A study of individual differences. *Cognitive Science*, *21*(1), 1-29.

Renkl, A. (1999). Learning mathematics from worked-out examples: Analyzing and fostering self-explanations. *European Journal of Psychology of Education, 14*(2), 477-488.

Renkl, A., & Atkinson, R. K. (2003). Structuring the transition from example study to problem solving in cognitive skill acquisition: A cognitive load perspective. *Educational Psychologist, 38*(1), 15-22.

Renkl, A., Stark, R., Gruber, H., & Mandl, H. (1998). Learning from worked-out examples: The effects of example variability and elicited self-explanations. *Contemporary Educational Psychology, 23*(2), 90-108.

Sliva, Y., Watson, G., & Morrison, G. R. (November, 2011). *Self-explanations and "expert-like" planning during problem-solving in geometry.* Paper presented at the 2011 International Association for Educational Communications and Technology (AECT) Convention, Jacksonville, FL.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257-285.

Sweller, J. (1999). *Instructional design in technical areas.* Camberwell, Australia: ACER Press.

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

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*(3), 185-233.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251-296.

Tennyson, R. D. (1980*).* Instructional control strategies and content structure as design variables in concept acquisition using computer-based instruction. *Journal of Educational Psychology, 72*(4), 525-532.

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

van Gerven, P. W. M., Paas, F., van Merriënboer, J. J. G., Hendriks, M., & Schmidt, H. G. (2003). The efficiency of multimedia learning into old age. *British Journal of Educational Psychology, 73*(4), 489–505.

van Gerven, P. W. M., Paas, F., van Merriënboer, J. J. G., & Schmidt, H. G. (2002). Cognitive load theory and aging: Effects of worked examples on training efficiency. *Learning and Instruction, 12*(1), 87–105.

van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist, 14*(1), 16-26.

van Merriënboer, J. J. G. (1990). Strategies for programming instruction in high school: Program completion vs. program generation. *Journal of Educational Computing Research, 6*(3), 265-287.

van Merriënboer, J. J. G., & Ayres, P. (2005). Research on cognitive load theory and it's design implications for e-learning, *Educational Technology Research and Development*, *53*(3), 5-13.

van Merriënboer, J. J.G., & de Croock, M. B. M. (1992). Strategies for computer-based programming instruction: Program completion vs. program generation. *Journal of Educational Computing Research*, *8*(3), 365-394.

van Merriënboer, J. J. G., Kester, L., & Paas F. (2006). Teaching complex rather than simple tasks: Balancing intrinsic and germane load to enhance transfer of learning. *Applied Cognitive Psychology*, *20*(3), 342-352.

van Merriënboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the load off a learner's mind: Instructional design for complex learning. *Educational Psychologist*, *38*(1), 5–13.

van Merriënboer, J. J. G., & Krammer, H. P. M. (1987). Instructional strategies and tactics for the design of introductory computer programming courses in high school. *InstructionalScience*, *16*(3), 251–285.

van Merriënboer, J. J. G, Schuurman, J. G., de Croock, M. B. M., & Paas, F. (2002). Redirecting learners' attention during training: Effects on cognitive load, transfer test performance and training. *Learning and Instruction*, *12*(1), 11–39.

van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, *17*(2),147–177.

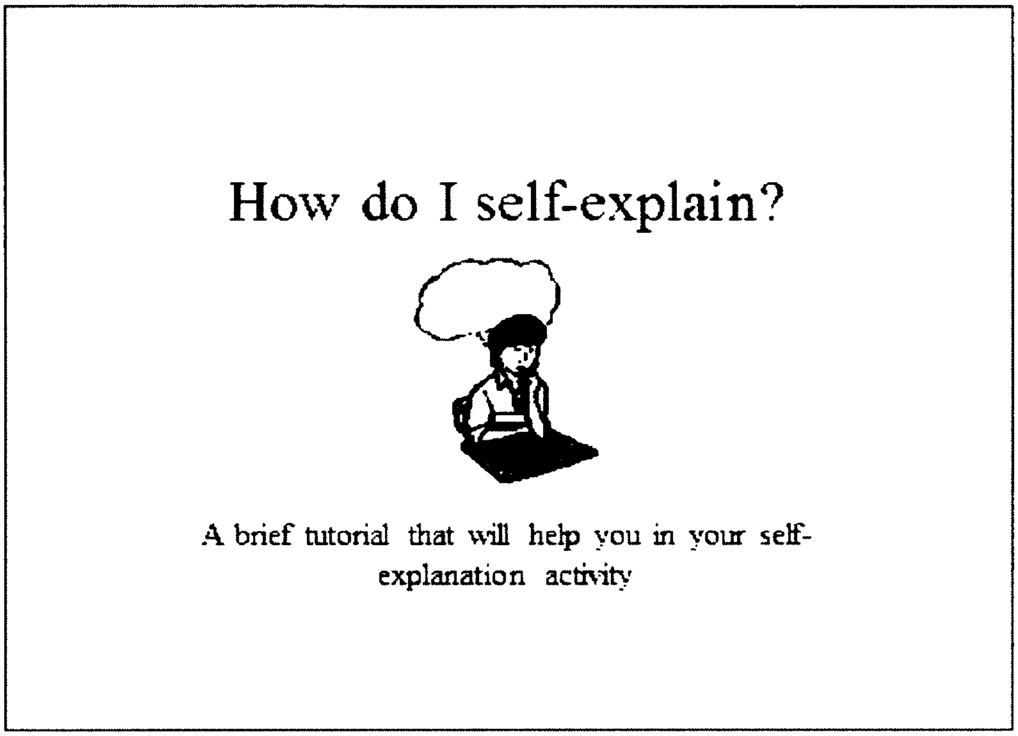
Yerkes, R.M., & Dodson, J.D. (1908). The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative Neurology and Psychology*, *18*, 459–482.

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology, 81*(3), 329-339.

Weinstein, C. E. (1982). Training students to use elaboration learning strategies. *Contemporary Educational Psychology, 7*(4), 301-311.

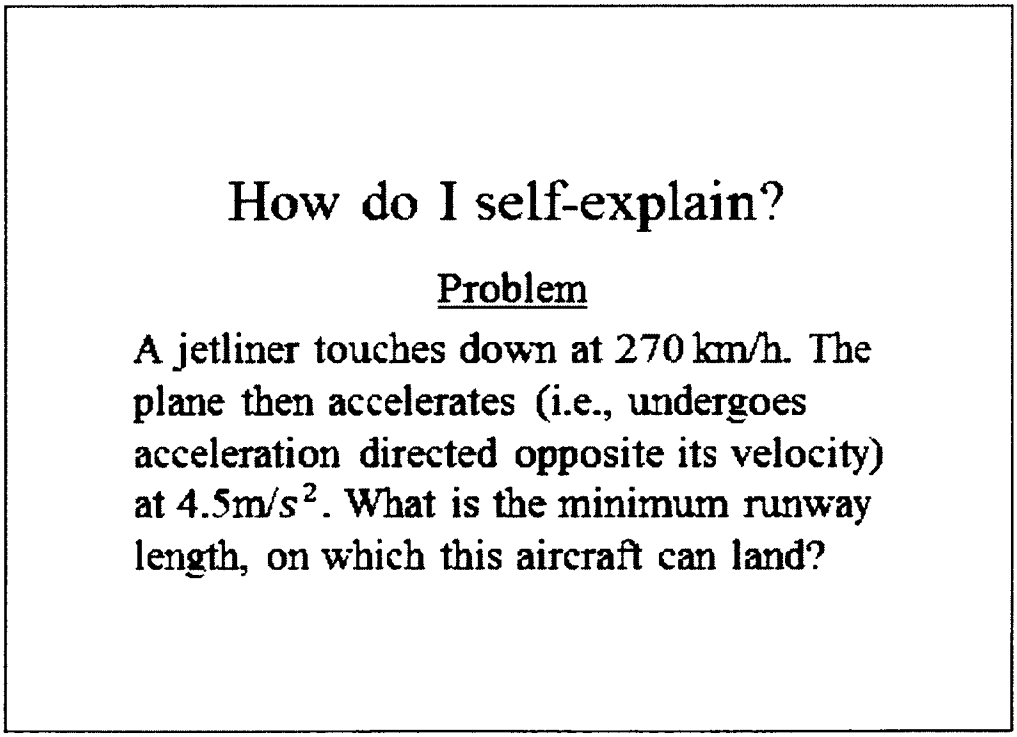
Wirth, J. (2009). Prompting self-regulated learning through prompts. *Zeitschrift für Pädagogische Psychologie, 23*(2), 91-94.

Appendix A. Tutorial on Self-Explanations



How do I self-explain?

A brief tutorial that will help you in your self-explanation activity



How do I self-explain?

Problem

A jetliner touches down at 270 km/h. The plane then accelerates (i.e., undergoes acceleration directed opposite its velocity) at 4.5m/s^2. What is the minimum runway length, on which this aircraft can land?

How do I self-explain?

Given:

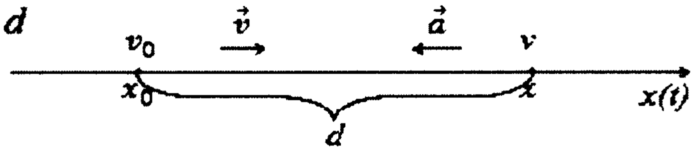
$v_0 = 270\,km/h$

$v = 0$

$a = -4.5m/s^2$

Find:

d



Step 1

Name the topic you are investigating

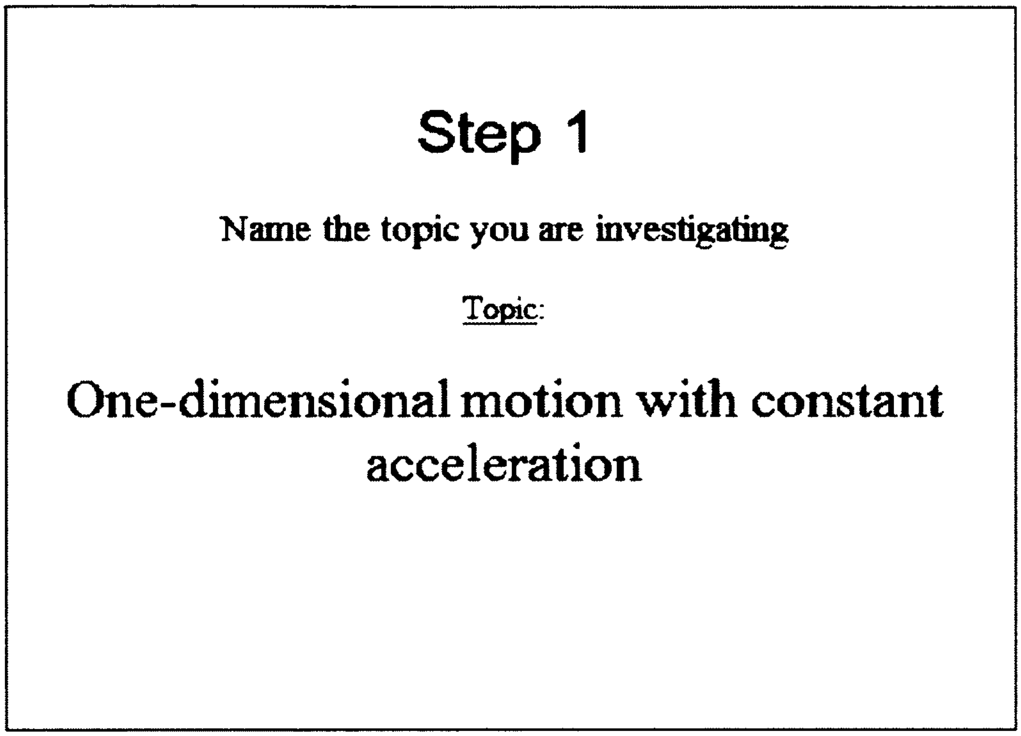
- Think about the particular situation you describe in the problem but in a much broader way

Is it a static situation or motion?
May be it is a collision?
What are the forces involved?
Or maybe you are asked about work and energy?
Or it could be a conservation of energy…

These are just examples of topics you go over in class.

- What you need to do is to classify the particular situation in a most general way

If you still have a problem with naming the topic, look into your recent lecture notes. The topics you discussed in class are those you have your homework assignment on. Look into worked example that relates to similar situation as in your homework problem and identify the section in your textbook where this example is located. What is the title of this section? It may be even in the title of the chapter.

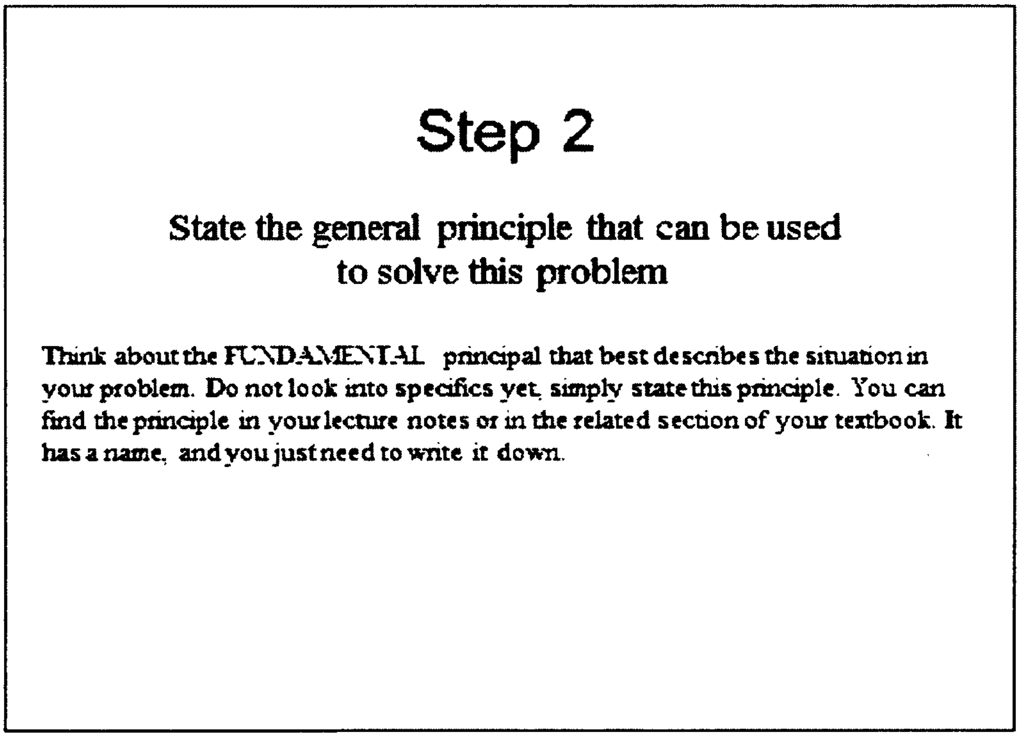


Step 1

Name the topic you are investigating

Topic:

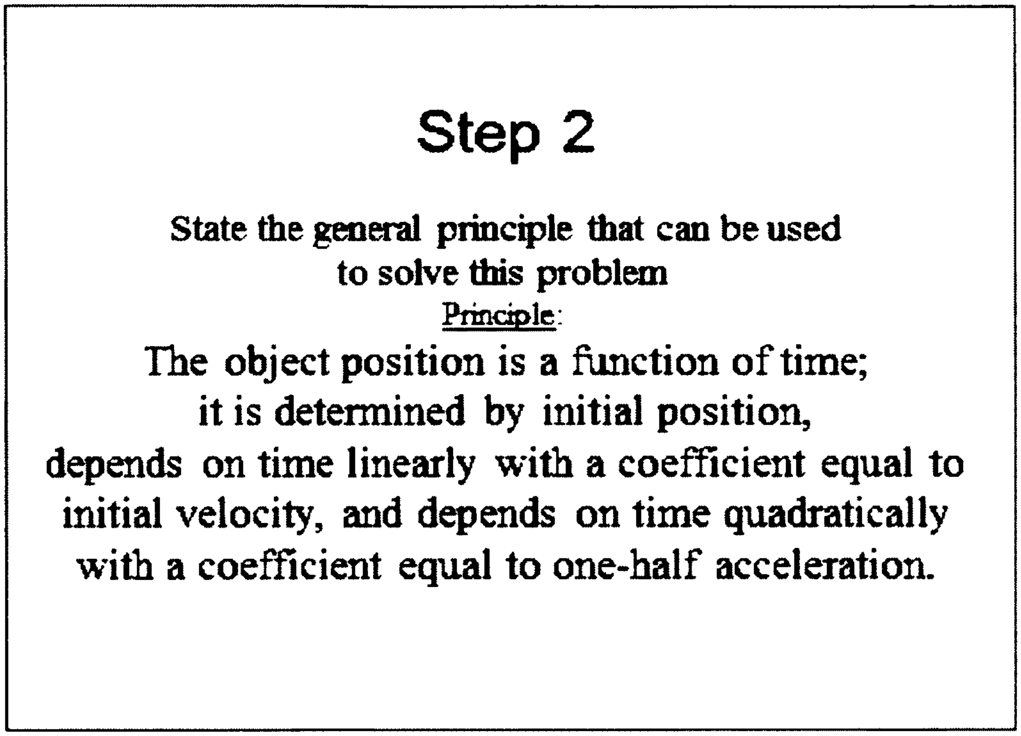
One-dimensional motion with constant acceleration



Step 2

State the general principle that can be used to solve this problem

Think about the FUNDAMENTAL principal that best describes the situation in your problem. Do not look into specifics yet, simply state this principle. You can find the principle in your lecture notes or in the related section of your textbook. It has a name, and you just need to write it down.

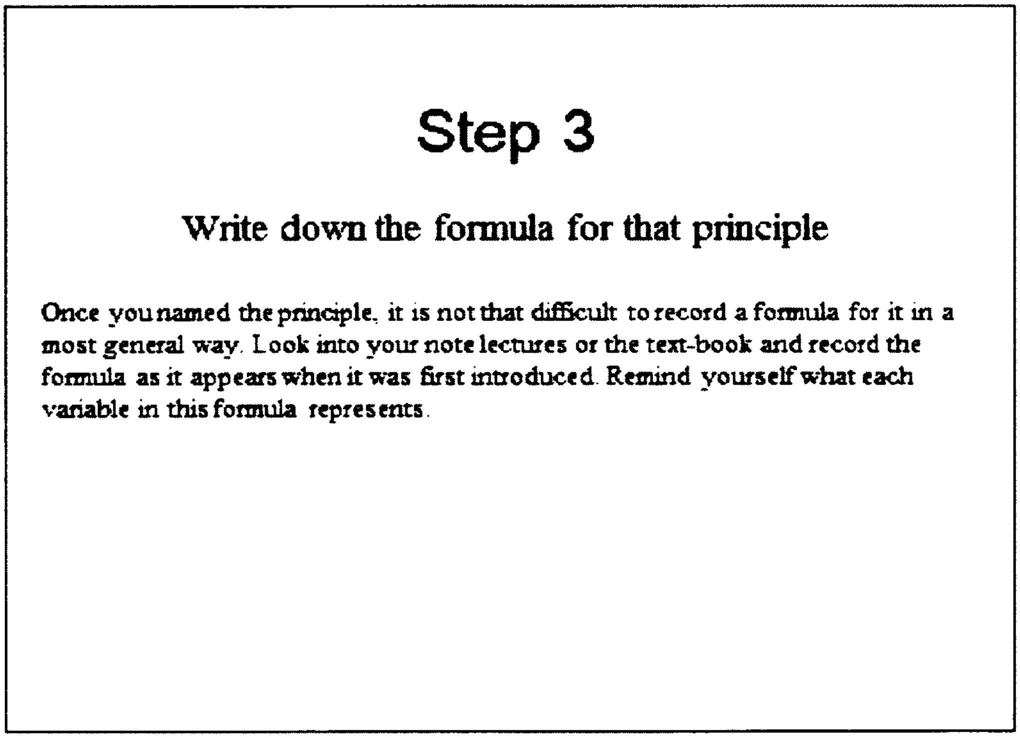


Step 2

State the general principle that can be used to solve this problem

Principle:

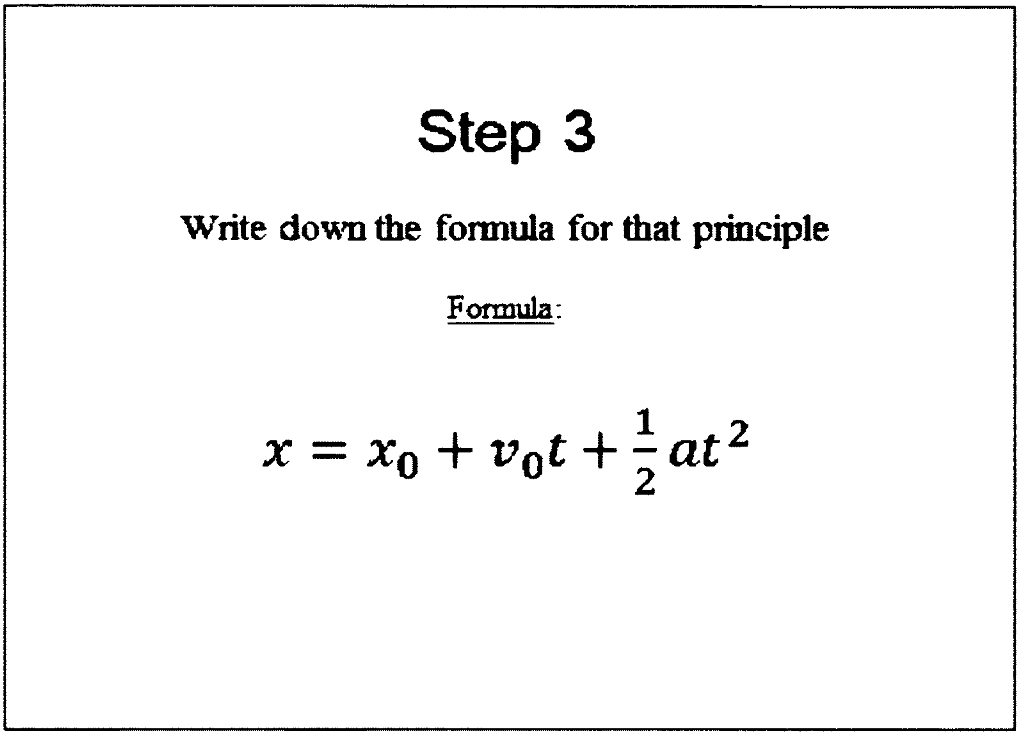
The object position is a function of time;
it is determined by initial position,
depends on time linearly with a coefficient equal to
initial velocity, and depends on time quadratically
with a coefficient equal to one-half acceleration.



Step 3

Write down the formula for that principle

Once you named the principle, it is not that difficult to record a formula for it in a most general way. Look into your note lectures or the text-book and record the formula as it appears when it was first introduced. Remind yourself what each variable in this formula represents.

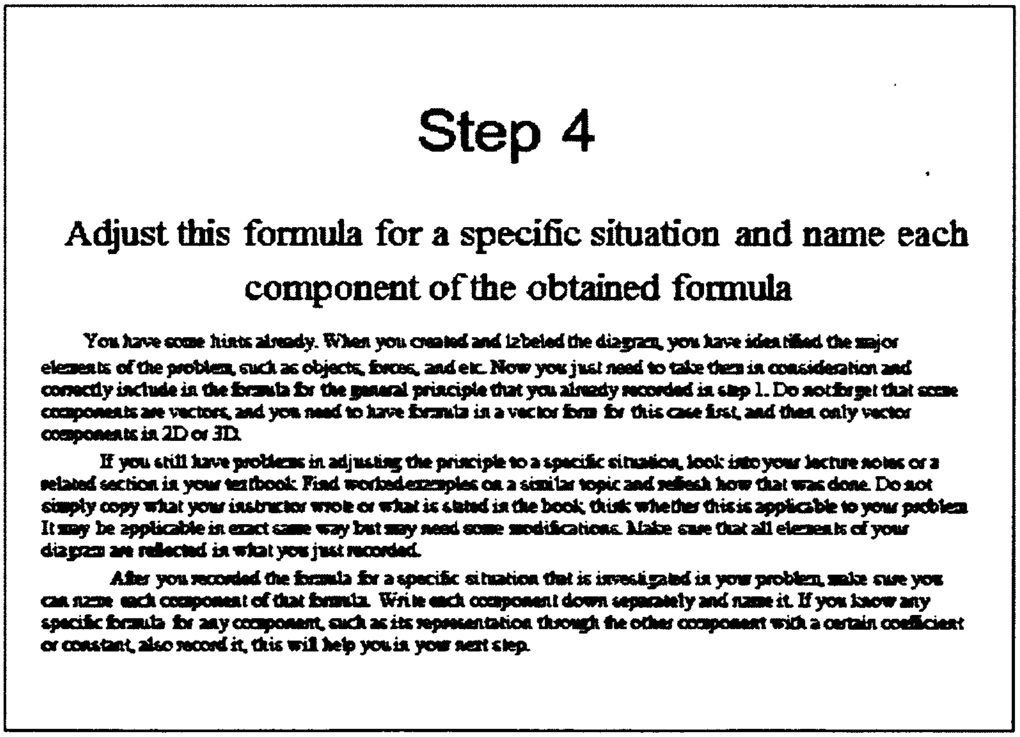


Step 3

Write down the formula for that principle

Formula:

$$x = x_0 + v_0 t + \frac{1}{2} a t^2$$



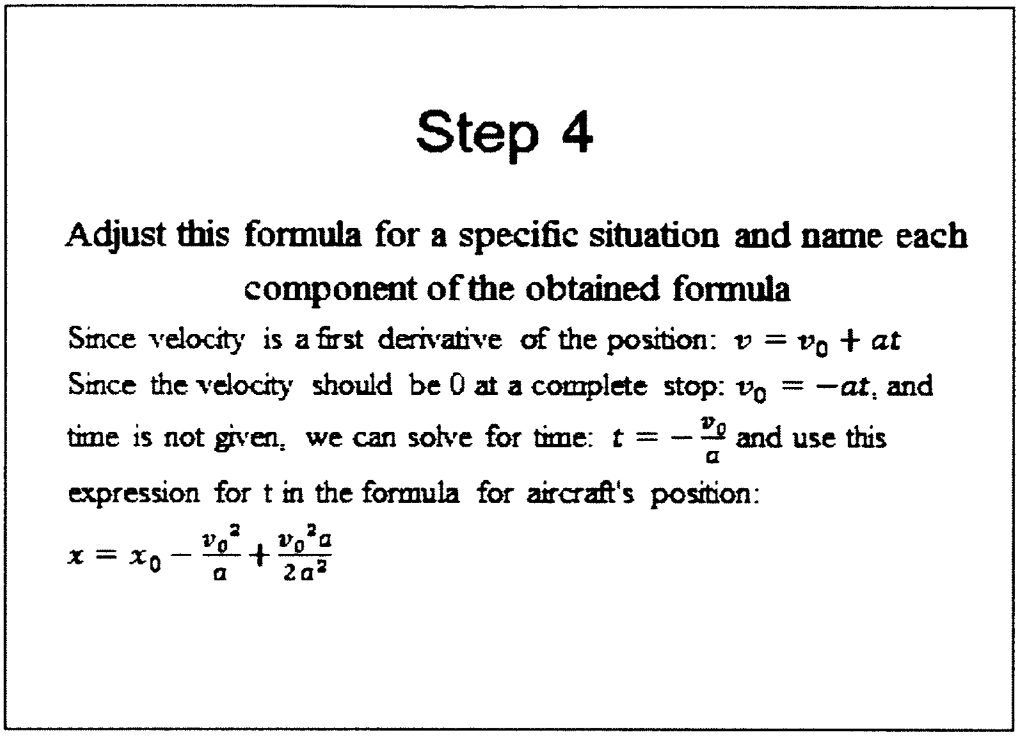
Step 4

Adjust this formula for a specific situation and name each component of the obtained formula

You have some hints already. When you created and labeled the diagram, you have identified the major elements of the problem, such as objects, forces, and etc. Now you just need to take them in consideration and correctly include in the formula for the general principle that you already recorded in step 1. Do not forget that some components are vectors, and you need to have formula in a vector form for this case first, and then only vector components in 2D or 3D.

If you still have problems in adjusting the principle to a specific situation, look into your lecture notes or a related section in your textbook. Find worked examples on a similar topic and refresh how that was done. Do not simply copy what your instructor wrote or what is stated in the book, think whether this is applicable to your problem. It may be applicable in exact same way but may need some modifications. Make sure that all elements of your diagram are reflected in what you just recorded.

After you recorded the formula for a specific situation that is investigated in your problem, make sure you can name each component of that formula. Write each component down separately and name it. If you know any specific formula for any component, such as its representation through the other component with a certain coefficient or constant, also record it, this will help you in your next step.

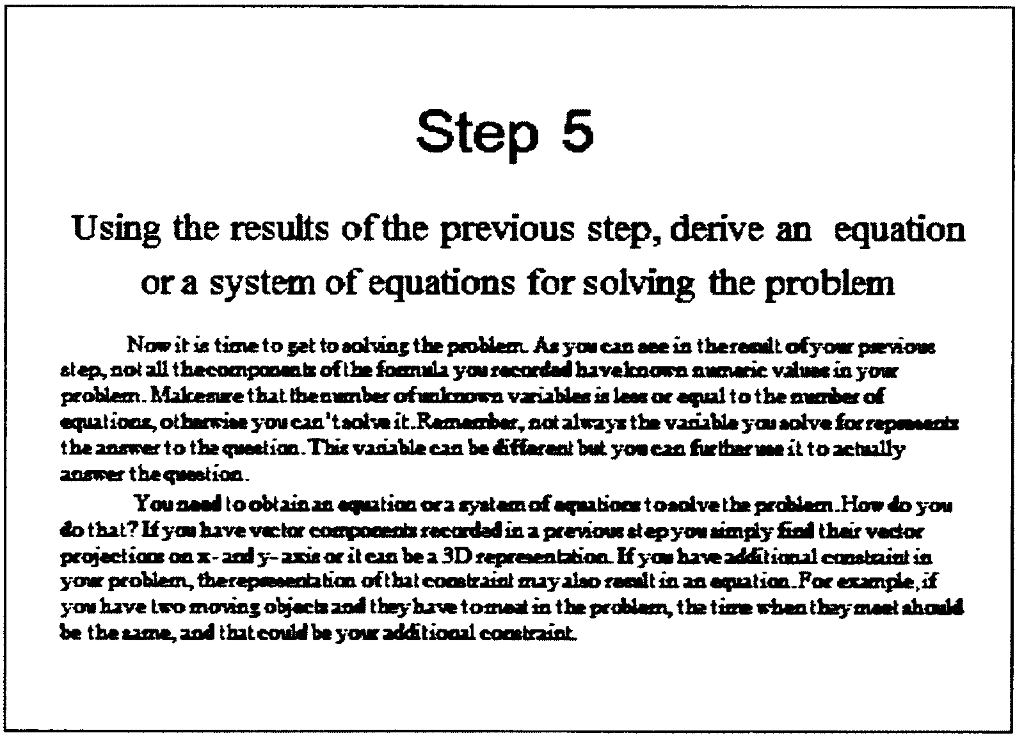


Step 4

Adjust this formula for a specific situation and name each component of the obtained formula

Since velocity is a first derivative of the position: $v = v_0 + at$
Since the velocity should be 0 at a complete stop: $v_0 = -at$, and time is not given, we can solve for time: $t = -\frac{v_0}{a}$ and use this expression for t in the formula for aircraft's position:

$$x = x_0 - \frac{{v_0}^2}{a} + \frac{{v_0}^2 a}{2a^2}$$



Step 5

Using the results of the previous step, derive an equation or a system of equations for solving the problem

Now it is time to get to solving the problem. As you can see in the result of your previous step, not all the components of the formula you recorded have known numeric values in your problem. Make sure that the number of unknown variables is less or equal to the number of equations, otherwise you can't solve it. Remember, not always the variable you solve for represents the answer to the question. This variable can be different but you can further use it to actually answer the question.

You need to obtain an equation or a system of equations to solve the problem. How do you do that? If you have vector components recorded in a previous step you simply find their vector projections on x- and y- axis or it can be a 3D representation. If you have additional constraint in your problem, the representation of that constraint may also result in an equation. For example, if you have two moving objects and they have to meet in the problem, the time when they meet should be the same, and that could be your additional constraint.

Step 5

Using the results of the previous step, derive an equation or a system of equations for solving the problem

Simplifying previously derived formula:

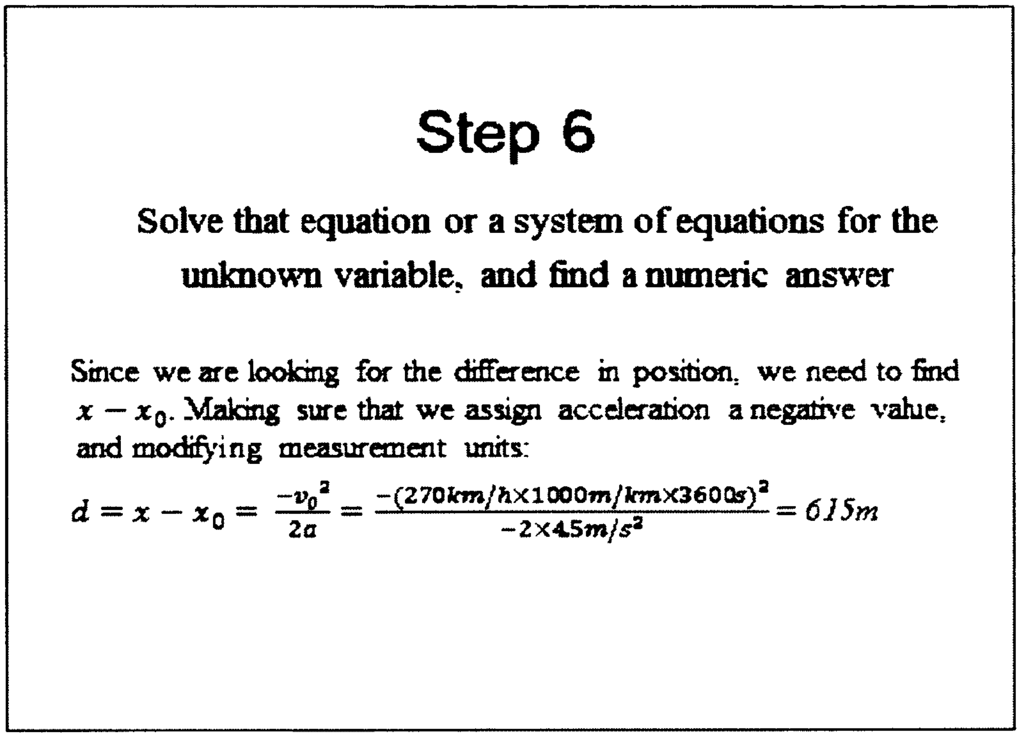
$$x = x_0 + \frac{-2v_0^2+v_0^2}{2a}$$

$$x = x_0 - \frac{v_0^2}{2a}$$

Step 6

Solve that equation or a system of equations for the unknown variable, and find a numeric answer

You are almost done. Make sure that you remember the two methods of algebra for solving systems of equations: elimination and substitution. Also remember that you have to keep your equations in terms of variables until the very last step, and then only substitute variables with their numeric values.

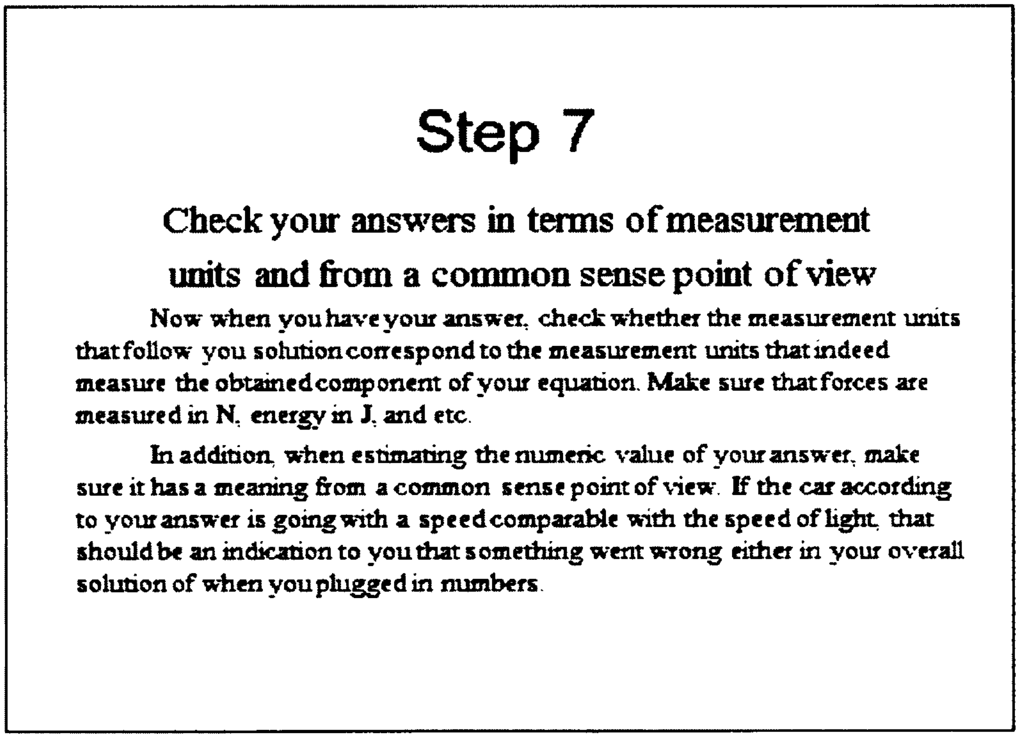


Step 6

Solve that equation or a system of equations for the unknown variable, and find a numeric answer

Since we are looking for the difference in position, we need to find $x - x_0$. Making sure that we assign acceleration a negative value, and modifying measurement units:

$$d = x - x_0 = \frac{-v_0^2}{2a} = \frac{-(270km/h \times 1000m/km \times 3600s)^2}{-2 \times 4.5m/s^2} = 615m$$

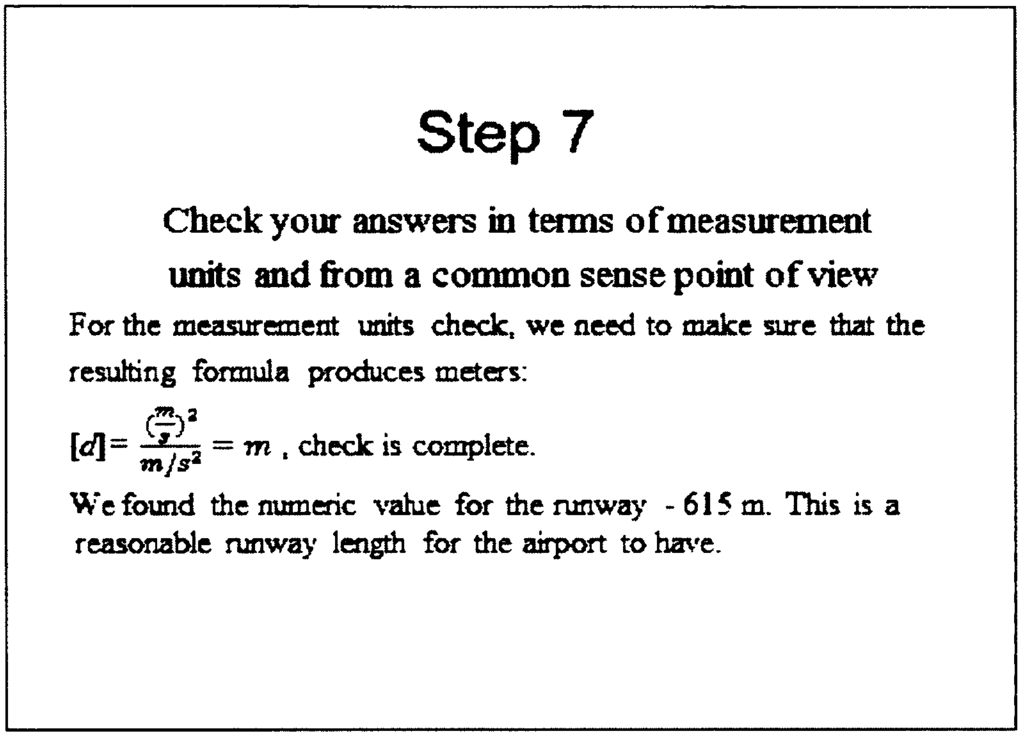


Step 7

Check your answers in terms of measurement units and from a common sense point of view

Now when you have your answer, check whether the measurement units that follow you solution correspond to the measurement units that indeed measure the obtained component of your equation. Make sure that forces are measured in N, energy in J, and etc.

In addition, when estimating the numeric value of your answer, make sure it has a meaning from a common sense point of view. If the car according to your answer is going with a speed comparable with the speed of light, that should be an indication to you that something went wrong either in your overall solution of when you plugged in numbers.



Step 7

Check your answers in terms of measurement units and from a common sense point of view

For the measurement units check, we need to make sure that the resulting formula produces meters:

$[d]=\frac{(\frac{m}{s})^2}{m/s^2}=m$, check is complete.

We found the numeric value for the runway - 615 m. This is a reasonable runway length for the airport to have.

Appendix B. Homework: Examples of Handouts

Control Group

Problem:

The 40 kg crate is positioned on a horizontal surface. The worker pulling a crate applies a force of 30 degrees to horizontal. If the coefficient of static friction between the crate and the surface is .650, with what force the worker must pull the crate for it to start moving.

Solution:

How much mental effort did you invest in solving this problem? (circle)

1--------2--------3--------4--------5--------6--------7--------8--------9

very,	**very,**
very low	**very high**

How much time did it take you to solve the problem?

Record the time you spent on this problem in _____ hours (if applicable) and _____ minutes

(do not include time on breaks or other not related activities)

Self-Explanation Group

Problem:

The 40 kg crate is positioned on a horizontal surface. The worker pulling a crate applies a force of 30 degrees to horizontal. If the coefficient of static friction between the crate and the surface is .650, with what force the worker must pull the crate for it to start moving.

Solution:

1. **Name the topic they are investigating**

2. **State the general principle that can be used to solve this problem**

3. **Write down the formula for that principle.**

4. **Adjust this formula for a specific situation and name each component of the obtained formula.**

5. **Using the results of the previous step, derive an equation or a system of equations for solving the problem.**

6. Solve that equation or a system of equations for the unknown variable, and find a numeric answer.

7. Check your answers in terms of measurement units and from a common sense point of view.

How much mental effort did you invest in solving this problem? (circle)

1--------2--------3--------4--------5--------6--------7--------8--------9

very, very low **very, very high**

How much time did it take you to solve the problem?

Record the time you spent on this problem in _____ hours (if applicable) and _____ minutes (do not include time on breaks or other not related activities)

Self-Explanation & Completion Group

Problem:

The 40 kg crate is positioned on a horizontal surface. The worker pulling a crate applies a force of 30 degrees to horizontal. If the coefficient of static friction between the crate and the surface is .650, with what force the worker must pull the crate for it to start moving.

Partial solution:

$\mathbf{F} + \mathbf{N} + \mathbf{W} + \mathbf{f} = \mathbf{0}$, where $f = \mu N$

Note: regular letters stand for the force magnitudes and bold letters stand for vector forces.

$F\cos\theta = \mu(mg - F\sin\theta)$

Solution:

1. **Name the topic they are investigating**

2. **State the general principle that can be used to solve this problem**

3. **Write down the formula for that principle.**

4. **Adjust this formula for a specific situation and name each component of the obtained formula.**

5. **Using the results of the previous step, derive an equation or a system of equations for solving the problem.**

6. **Solve that equation or a system of equations for the unknown variable, and find a numeric answer.**

7. **Check your answers in terms of measurement units and from a common sense point of view.**

How much mental effort did you invest in solving this problem? (circle)

1--------2--------3--------4--------5--------6--------7--------8--------9

very, very low **very, very high**

How much time did it take you to solve the problem?

Record the time you spent on this problem in _____ hours (if applicable) and _____ minutes

(do not include time on breaks or other not related activities)

Appendix C. Sample Homework Submissions

Problem:

The 40 kg crate is positioned on a horizontal surface. The worker pulling a crate applies a force of 30 degrees to horizontal. If the coefficient of static friction between the crate and the surface is .650, with what force the worker must pull the crate for it to start moving.

***Control group** will be provided with the following sample submission:*

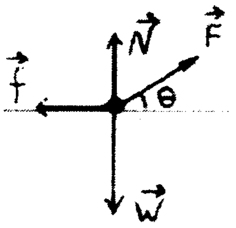
Given: $m = 40.0$ kg

$\theta = 30°$

$g = 9.8\ m/s^2$

$\mu = .650$

Find: F



$\mathbf{F + N + W + f = 0}$

x-axis: $F\cos\theta - \mu N = 0$ (i)

y-axis: $N + F\sin\theta - mg = 0$ (ii)

$F\cos\theta = \mu N$ (i)

$N = mg - F\sin\theta$ (ii)

Combining (i) and (ii):

$F\cos\theta = \mu(mg - F\sin\theta)$

Solving for F:

Fcosθ = μmg - μFsinθ

Fcosθ + μFsinθ = μmg

F(cosθ + μsinθ) = μmg

$$F = \frac{\mu mg}{\cos\theta + \mu\sin\theta} = \frac{0.650 \cdot 40.0kg \cdot 9.80\frac{m}{s^2}}{0.866 + 0.650 \cdot 0.5} \approx 214N$$

Self-explanation group *will be given the sample submission*:

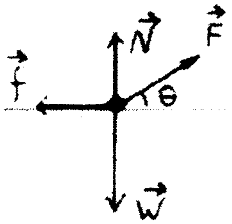
Given: m = 40.0 kg

θ = 30°

g = 9.8 m/s²

μ = .650

Find: F



1. Topic: The force is acting on the object; motion with friction; applied force has to slightly exceed the maximum static friction force.
2. Principle: Newton's second law of motion in case of equilibrium
3. Σ **F** = **0**
4. Principle applied to a specific situation:

F + **N** + **W** + **f** =**0** (all vector forces acting on the crate), where

F applied force at an angle θ to horizontal;

N normal to the surface force;

W = mg gravitational force;

and **f** is a kinetic friction force acting in the direction opposite to potential moving and magnitude $f = \mu N$

5. Obtaining a system of equations:

   vector projection of equation obtained in step 1 on x-axis: $F\cos\theta - \mu N = 0$ (i)

   vector projection of equation obtained in step 1 on y-axis: $N + F\sin\theta - mg = 0$ (ii)

6. Solving (algebraically and numerically) system of equations to find F:

   from (i): $F\cos\theta = \mu N$

   from (ii): $N = mg - F\sin\theta$

   combining (i) and (ii):

   $F\cos\theta = \mu(mg - F\sin\theta)$ (iii)

   solving (iii):

   $F\cos\theta = \mu mg - \mu F\sin\theta$

   $F\cos\theta + \mu F\sin\theta = \mu mg$

   $F(\cos\theta + \mu\sin\theta) = \mu mg$

$$F = \frac{\mu mg}{\cos\theta + \mu\sin\theta} = \frac{0.650 \cdot 40.0kg \cdot 9.80\frac{m}{s^2}}{0.866 + 0.650 \cdot 0.5} \approx 214N$$

7. Checking measurement units:

   $(kg \cdot m)/s^2 = N$ (makes sense)

Self-explanation & completion group *will be given the sample submission:*

Note, partial solution is highlighted gray.

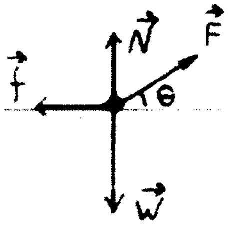
Given: $m = 40.0$ kg

$\theta = 30°$

$g = 9.8$ m/s^2

$\mu = .650$

Find: F



1. Topic: The force is acting on the object; motion with friction; applied force has to slightly exceed the maximum static friction force.
2. Principle: Newton's second law of motion in case of equilibrium
3. $\Sigma \mathbf{F} = \mathbf{0}$
4. Principle applied to a specific situation:

   [redacted] (all vector forces acting on the crate), where

   F applied force at an angle θ to horizontal;

   N normal to the surface force;

   W = mg gravitational force;

   and **f** is a kinetic friction force acting in the direction opposite to potential moving and magnitude $f = \mu N$

5. Obtaining a system of equations:

   vector projection of equation obtained in step 1 on x-axis: $F\cos\theta$ - [redacted] = 0 (i)

   vector projection of equation obtained in step 1 on y-axis: $N + F\sin\theta - mg = 0$ (ii)

6. Solving (algebraically and numerically) system of equations to find F:

   from (i): $F\cos\theta = \mu N$

   from (ii): $N = mg - F\sin\theta$

   combining (i) and (ii):

[redacted] (iii)

solving (iii):

$F\cos\theta = \mu mg - \mu F\sin\theta$

$F\cos\theta + \mu F\sin\theta = \mu mg$

$F(\cos\theta + \mu\sin\theta) = \mu mg$

$$F = \frac{\mu mg}{\cos\theta + \mu\sin\theta} = \frac{0.650 \cdot 40.0kg \cdot 9.80\frac{m}{s^2}}{0.866 + 0.650 \cdot 0.5} \approx 214N$$

7. Checking measurement units:

(kg·m)/s² = N (makes sense)

Appendix D. Partial Credit Criteria and Related Points for Near and Far Transfer Quiz and Posttest Problems

Criteria	Points
Correct algebraic solution, correct numeric answer	5
Correct algebraic solution, calculation error	4
Correct approach to the problem, however, algebraic solution has minor errors or student did not finish solving the problem (no final algebraic solution produced)	3
Student made an attempt to solve the problem; wrote relevant formulas but failed to assign them correct context-related meaning	2
Student did not make an attempt to solve the problem, only wrote a couple of general formulas	1
Student left the paper blank	0

Appendix E. Story Problem Scoring Rubric

Item	Criteria			Points
	fully accomplished	partially accomplished	not attempted/ response is meaningless	
Did the student provide and accurately label the diagram for the problem?	2	1	0	
Did the student provide a correct formula for the principle to solve the problem?	2	1	0	
Did the student derive an equation or a system of equations that follows the correct application of that principle to a specific situation?	2	1	0	
Did the student correctly solve the problem algebraically?	2	1	0	
Did the student obtain a correct numeric answer to the problem?	2	1	0	
			Total	

Appendix F. Summary of Posttest Grading

Rote level problems			**Problems (combined)**				**Story problems**				**Post-test**
	max. points possible		Sub-questions	max. points possible		total point possible		max. points possible		total points possible	post test points possible
Recall	1	x 4									4
			a) Comprehension	2		8					8
			b) Near transfer	5							
			c) Near transfer	5							
			Total near transfer	10	x 4	40					40
			d) Far transfer	5							
			e) Far transfer	5							
			Total far transfer	10		40	Far transfer	10	x 3	30	70
Problem total	1		Problem total	22			Problem total	10			
										Post test total	122

Appendix G. Quality of Self-Explanations Score Rubric

Did the student name the investigated topic?			
Named correctly-3	Named partially correct -2	Named but incorrectly-1	Did not name-0
Did the student name the general principle for solving the problem?			
Named correctly-3	Named partially correct -2	Named but incorrectly-1	Did not name-0
Did the student adjust that principle to a specific situation?			
Adjusted correctly-3	Adjusted but was partially correct-2	Attempted to adjust but the response is meaningless-2	Did not attempt-0
Did the student explain each component of the obtained formula for the application of the principle to a specific situation?			
Explained all components correctly -3	Explained most of components correctly-2	Explained but most of them incorrectly-1	Did not explain or explained but all incorrectly-0
Did the student derive an equation or a system of equations?			
Derived correctly-3	Derived but was partially correct-2	Attempted to derive but the response is meaningless-1	Did not attempt-0
Did the student solve the obtained equation or a system of equations?			
Solved correctly-3	Solved but was partially correct-2	Attempted to solve but the response is meaningless-1	Did not attempt-0
Did the student check and confirm the correctness of measurement units?			
Checked and confirmed they were correct-3	Although checked, was not able to confirm-2	Although checked and confirmed, measurement units make no sense-1	Did not check-0

Appendix H. Survey

Distributed to the participants at the end of the study
"In Search for Instructional Techniques to Maximize the Use of Germane Cognitive Resources: A Case of Teaching Complex Tasks in Physics"

Please circle the responses that are best applicable to you for each of the following statements

1. In this study I was assigned to the following group:

A B C

For the statements below, please use the following rating scale:

2. During the study I did my homework in a same way I did it before the study.

1	2	3	4	5	6	7
Strongly agree	Agree	Slightly agree	Neither agree, nor disagree	Slightly disagree	Disagree	Strongly disagree

3. The way I approached my homework made me think about the problems in-depth.

1	2	3	4	5	6	7
Strongly agree	Agree	Slightly agree	Neither agree, nor disagree	Slightly disagree	Disagree	Strongly disagree

4. I do not think that the way I did my homework during this study helped me to better prepare for quizzes and the test.

1	2	3	4	5	6	7
Strongly agree	Agree	Slightly agree	Neither agree, nor disagree	Slightly disagree	Disagree	Strongly disagree

5. As I gained experience, I changed my approach to solving problems and implemented it during quizzes and the test.

1	2	3	4	5	6	7
Strongly agree	Agree	Slightly agree	Neither agree, nor disagree	Slightly disagree	Disagree	Strongly disagree

If you changed your approach, briefly explain what change(s) you made:

____

____

6. After the study is over, I will not continue doing my homework the way I was prompted during the study.

1	2	3	4	5	6	7
Strongly agree	Agree	Slightly agree	Neither agree, nor disagree	Slightly disagree	Disagree	Strongly disagree

7. The approach to solving homework problems introduced in this study should be used for teaching physics.

1	2	3	4	5	6	7
Strongly agree	Agree	Slightly agree	Neither agree, nor disagree	Slightly disagree	Disagree	Strongly disagree

8. Overall, I had positive experience with the instructional method introduced to me in this study.

7	6	5	4	3	2	1
Strongly agree	Disagree	Slightly disagree	Neither agree, nor disagree	Slightly agree	Agree	Strongly disagree

9. Overall, I think I benefited from participating in this study.

1	2	3	4	5	6	7
Strongly agree	Agree	Slightly agree	Neither agree, nor disagree	Slightly disagree	Disagree	Strongly disagree

Your comments on the next six survey items are greatly appreciated:

10. Please name one aspect you especially liked about the introduced approach to solving problems

____

____

11. Please name one aspect you disliked about the introduced approach

____

____

12. Please explain how the approach introduced in this study helped you with learning physics. If you think it didn't help you, please circle "It didn't help" and give your reasons.

____

____

____

____

____

13. If after the study you plan to continue solving physics problems the way that was introduced to you, please explain why you would continue using the suggested approach. If you do not plan to continue, please circle "I don't plan" and give your reason.

14. If you think the approach introduced in this study should be used for teaching physics please explain why. If you don't think it should be used, please circle "Shouldn't be used" and give your reason.

15. Any other comments, thoughts, or suggestions you want to share about the suggested approach to solving problems or the study.

Appendix I. Informed Consent Form

PROJECT TITLE:

In Search for Instructional Techniques to Maximize the Use of Germane Cognitive Resources: A Case of Teaching Complex Tasks in Physics

RESEARCHERS

Dr. Gary R. Morrison, Professor, **Responsible Project Investigator**
Old Dominion University, Darden College of Education, STEM Education and Professional Studies
gmorriso@odu.edu

Dr. Ginger S. Watson, Associate Professor
Old Dominion University, Darden College of Education, STEM Education and Professional Studies
gswatson@odu.edu

Dr. Alexander L. Godunov, Associate Professor, College of Sciences, Department of Physics
agodunov@odu.edu

Yekaterina Sliva, Doctoral Student
Old Dominion University, Darden College of Education, STEM Education and Professional Studies
ysliv001@odu.edu

PURPOSE

The purposes of this consent form are to provide you information that may affect your decision whether to participate in this research, and to record the consent of those who agree to participate. Multiple studies have investigated problem-solving and transfer of knowledge in complex domains of knowledge, such as physics. However, most studies were short-term and focused on one specific strategy applied to one selected topic. This study will investigate an instructional technique that can be used throughout a semester-long course and may be applicable to any topic of this course, as well as to other subjects that require learning of complex content.

If you decide to participate, you will be asked to complete your regular homework assignments in a specific format each week for four weeks. You will receive specific instructions on the format of your submissions via a Blackboard tutorial prior to the study. You will also complete four weekly quizzes and one posttest at the end of the study. Completion of your homework assignments, quizzes, and tests are required activities for this course, regardless of whether you participate in this study or not. If you participate in this study you'll be asked to report your mental effort and record time on task for each problem. To rate invested mental effort, you will be specifically asked: "How much mental effort did you invest in solving this problem?" and you will rate your responses from1 (very, very low) to 9 (very, very high). You will be asked to provide the same rating for each question on four quizzes and one test. To record time on task, you will be asked "How much time did it take you to solve the problem?" and you will record the time you spent on the problem in hours (if applicable) and minutes. Recording your mental effort and time to complete each problem will take about 10 minutes for each homework assignments, one minute for quizzes, and five minutes for the test. At the end of the study, you will be asked to respond to a short attitude survey during the regular class meeting, which will take you about 15 minutes. No additional time (outside of the regularly scheduled class session) will be required of you for any of the activities associated with this research.

RISKS AND BENEFITS

RISKS: If you decide to participate in this study, then you may face a risk of breach of confidentiality. However, the researchers will try to reduce this risk by replacing your name with a unique study code to link your homework, quiz, and test data for this study. Yet, with any research, there is a possibility that you may be subject to risks that have not yet been identified.

BENEFITS: There will be no direct benefits offered for participation in this study. The information gained from this study will add to our understanding of cognitive processes taking place in learner's mind while they are engaged in active information processing. You and your classmates who chose to participate in this research will benefit from the study by being introduced to instructional techniques that may help you to improve your problem-solving skills and better prepare for tests.

COSTS AND PAYMENTS

Participation in this study will not involve any additional costs to you. The researchers are unable to give you any payments for participating in this study, however for those who agree to participate and complete all required work, a drawing of 20 raffle tickets for amazon.com gift certificates in the amount of $50 each will take place at the end of the study.

NEW INFORMATION

If the researchers find new information during this study that would reasonably change your decision about participating, then they will give it to you.

CONFIDENTIALITY

The researchers will take reasonable steps to keep private information confidential. The researchers will use a unique study code to identify each participant. Student names, initials, University Identification Numbers (UIN), and personally identifiable information will **not** be used. During data collection, one researcher will keep a master list of student names and their codes to ensure that each student's data are linked between weeks. This list will be kept in a locked file and only the researchers will have access to the list. The master list linking student names to their study code will be discarded when the study data collection and analysis are complete. The researchers will not share identifiable information collected during this study with anyone outside of the research team. The results of this study may be used in reports, presentations, and publications. However, the researchers will not identify you. Of course, your records may be subpoenaed by court order or inspected by government bodies with oversight authority.

WITHDRAWAL PRIVILEGE

It is OK for you to say NO. Even if you agree to participate now, you are free to walk away or withdraw from the study at any time. Your decision will not affect your relationship with your instructor or university. The researchers reserve the right to withdraw your participation in this study, at any time, if they observe potential problems with your continued participation.

COMPENSATION FOR ILLNESS AND INJURY

If you agree to participate, your consent in this document does not waive any of your legal rights. However, in the event of harm arising from this study, neither Old Dominion University nor the researchers are able to give you any money, insurance coverage, free medical care, or any other compensation for such injury. In the event that you suffer injury as a result of participation in any research project, you may contact Dr. Gary R. Morrison (gmorriso@odu.edu) or 757-683-6275, who will be glad to review the matter with you.

VOLUNTARY CONSENT

By agreeing to be a part of this study, you are saying several things. You are saying that you have read this form or have had it read to you, that you have understood this form, the research study, and its risks and benefits. Feel free to keep a copy of this form for your records. The researchers should have answered any questions you may have had about the research. If you have any additional questions later on, then you can contact Dr. Gary R. Morrison (gmorriso@odu.edu or 757-683-6275). If at any time you feel pressured to participate, or if you have any questions about your rights or this form, then you should contact Dr. George Maihafer the current IRB chair at 757-683 4520 at Old Dominion University, or the Old Dominion University Office of Research at 757-683-3460.

If you do not agree to be a part of this study, then simply return this form to the researcher. If you agree to be a part of this study, then put your name, signature and today's date on the line below.

Participant'sname___________________Signature_______________Date_________

INVESTIGATOR'S STATEMENT

I certify that I have explained to this subject the nature and purpose of this research, including benefits, risks, costs, and any experimental procedures. I have described the rights and protections afforded to human subjects and have done nothing to pressure, coerce, or falsely entice this subject into participating. I am aware of my obligations under state and federal laws, and promise compliance. I have answered the subject's questions and have encouraged him/her to ask additional questions at any time during the course of this study.

Investigator's Printed Name & Signature	**Date**

VITA

Yekaterina Sliva

Old Dominion University
STEM and Professional Studies, Darden College of Education
Norfolk, Virginia

Education

Old Dominion University (Darden College of Education) *| Norfolk, VA*
December 2013

- PhD in Instructional Design & Technology Program from the Department of STEM Education and Professional Studies

Moscow State University *| Moscow, Russia*
January 1990

- Combined B.S. and M.S. in Physics

Professional Experience

Old Dominion University, Lecturer of Mathematics *| Norfolk, VA*
07/2007-05/2010

- Given course objectives, developed curriculum and taught the following undergraduate level courses: Developmental Mathematics, PreCalculus & Trigonometry, Geometry, Calculus I, and Calculus II
- Member of the UExcel Examination Development Committee in Calculus (www.uexceltest.com)
- Served as a member of a committee that was set up to oversee the merger of the Developmental Mathematics Center with the Department of Mathematics and Statistics
- Screened prospective MS students for the Teacher Immersion Residency Program as delegated by the Teacher Quality Partnership Grant received by ODU Research Foundation from the U.S. Department of Education

Old Dominion University, Mathematics Instructor *| Norfolk, VA*
07/2006-07/2007

- Given course objectives, developed curricula and taught the following undergraduate level courses: Mathematics for Liberal Arts, Algebra, PreCalculus

Old Dominion University, Adjunct Mathematics Instructor *| Norfolk, VA*
09/2005-07/2006

- Taught the following undergraduate courses: Mathematics for Liberal Arts and Algebra

Thomas Nelson Community College, Adjunct Mathematics Instructor *| Hampton, VA*
01/2005-05/2005

- Taught the following undergraduate courses: Consumer Math, PreAlgebra, Algebra, PreCalculus, Probability & Statistics

Phoebus High School, Mathematics Teacher *| Norfolk, VA*
09/2004-01/2005

- Taught Algebra primarily to 9th grade students, as well as students who had repeatedly failed Algebra

Tidewater Tech (currently Centura College), Adjunct Mathematics Instructor *| Newport News, VA*
02/2004-06/2004

- Taught Consumer Mathematics and PreAlgebra/Algebra
- Due institutional restructuring, developed a new mathematics curriculum according to state accreditation requirements

Academic Honors

09/2012-05/2013

- Dissertation Fellowship, ODU Darden College of Education

Publications/Presentations

Sliva, Y., Watson, G., & Morrison, G. (November, 2011). *Self-explanations and "expert-like" planning during problem-solving in geometry.* Paper presented at the 2011 International Association for Educational Communications and Technology (AECT) Convention, Jacksonville, FL.

Romero, E., Sliva, Y., Brown, J., & Watson, G. (November, 2011). *Optics reflection and refraction: A virtual instructional simulation.* Paper presented at the 2011 International Association for Educational Communications and Technology (AECT) Convention, Jacksonville, FL.

Sliva, Y. (2013, April). *In Search of Instructional Techniques to Maximize the Use of Germane Cognitive Load: A Case of Teaching Complex Tasks in Physics.* Paper presented at the Old Dominion University Annual Graduate Research Achievement Day, Norfolk, VA.

Professional Affiliations

Association for Educational Communications and Technology (AECT)