The Pennsylvania State University

The Graduate School

College of Education

THE EFFECTS OF SEGMENTED MULTIMEDIA WORKED EXAMPLES AND SELF-EXPLANATIONS ON ACQUISITION OF CONCEPTUAL KNOWLEDGE AND PROBLEM-SOLVING PERFORMANCE IN AN UNDERGRADUATE ENGINEERING COURSE

A Dissertation in

Instructional Systems

by

Natalia V. Kapli

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2010

The dissertation of Natalia V. Kapli was reviewed and approved* by the following:

Barbara L. Grabowski
Professor of Education
Dissertation Advisor
Chair of Committee

Susan M. Land
Associate Professor of Education

Peggy N. Van Meter
Associate Professor of Education

David L. Passmore
Professor of Education

Simon R. Hooper
Associate Professor of Education
In Charge of Graduate Programs in Instructional Systems

*Signatures are on file in the Graduate School

ABSTRACT

The study investigated the effects of non-segmented multimedia worked examples (NS-MWE), segmented multimedia worked examples (S-MWE), and segmented multimedia worked examples enhanced with self-explanation prompts (S-MWE-SE) on acquisition of conceptual knowledge and problem solving performance in an undergraduate engineering course. In addition, the study examined learner-generated self-explanations for possible interrelationships between the type and quality of self-explanations and learning outcomes.

A total of 62 engineering students from a 300-level civil engineering course participated in the study as a part of their course. After the initial introduction of the concepts by the instructor in class, they completed the treatment and the conceptual knowledge test online as a part of their homework assignment, and completed a problem-solving test in class as a part of a course quiz. Participants' homework logs were retrieved to examine the generated self-explanations, and determine time on task and amount of access to the multimedia files.

A one-way ANOVA test did not reveal any significant differences between any of the groups on either outcome measure. Time on task also did not reveal any significant differences between the groups. However, in-treatment problem solving performance scores showed a significant difference between the S-MWE-SE and the NS-MWE groups with the S-MWE-SE group performing significantly better.

There were several positive linear relationships found between the principle-based type of explaining, quality of explaining, and problem solving performance, as well as between correctness of self-explanations and acquisition of conceptual knowledge. Multiple regression results revealed that quality of self-explanations was a significant predictor of performance on the problem-solving test for the S-MWE-WE group.

Multimedia access logs from the groups that received segmented multimedia example treatments were examined for possible patterns using a two-step clustering procedure. As a result, three distinct groups were identified based on their amount of access to each of the 7 segments: the most frequent viewers, moderate viewers, and the least frequent viewers. One-way ANOVA revealed that the most frequent viewers performed significantly worse on the problem-solving test than the other two groups. The results followed a similar pattern on the conceptual knowledge test score, but the differences did not reach the significance level. However, all three types of viewers received very similar scores on the in-treatment performance.

It was also discovered from the logs that several participants may have collaborated on the treatment. Both collaborating and non-collaborating groups were examined to determine whether any possible differences in performance existed. There was a significant difference in performance on the conceptual knowledge test with the collaborating group performing significantly better. Multiple regression results revealed that principle-based explaining was a significant predictor of performance on the problem-solving test for non-collaborating participants in the S-MWE-WE group.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deep appreciation to Dr. Barbara Grabowski, my academic and thesis advisor, for her guidance and encouragement through all these years, as well as for her outstanding editing expertise. I learned what it means to be a scholar and a mentor from her. I would like to thank my committee members for their guidance and advice on various aspects of the research process. My sincere thanks go to Dr. Peggy Van Meter for her invaluable insights on the research design. Her class on conceptual learning and problem solving helped me shape the theoretical framework for this study. I would like to thank Dr. David Passmore for his advice on statistical procedures, his encouragement and sense of humor. I would also like to thank Dr. Susan Land for bringing a different perspective to the table, for her continuous support, and for giving me a hug right before my defense, I really needed it.

This study would not have been possible without the help of the Leonhard Center for the Enhancement of Engineering Education and its director Dr. Thomas Litzinger, who gave me access to the College of Engineering faculty, and who also took on the role of one of my subject matter experts at one of the most important phases of the study. I would also like to express my deep appreciation to Dr. Roxanne Toto, who spent countless hours planning and implementing this project, being both a technical expert and a supportive colleague.

I would like to express my sincere gratitude to Dr. Michael Gooseff who generously allowed me to use his class as a research setting and provided the richest multimedia content I could ever wish for. I could never thank him enough for willing to make changes to his course to accommodate this study. I would also like to thank Dr. Gooseff’s graduate assistant, Tom Voltz, for his incredible patience in explaining to me how to solve fluid mechanics problems while we were working on developing a scoring rubric.

I attribute a part of my success to Dr. Vicki Williams, my work supervisor at Education Technology Services. She created a very positive working environment for me and was very supportive and understanding throughout the process.

I would like to thank my friends in the Instructional Systems program, especially Dr. Grabowski's research apprenticeship group, for being able to share my ideas with them and receive valuable feedback over all these years.

I would like to thank my parents, Tatiana and Vladimir Kapli, for teaching me to always strive for the best in everything I do. They have been waiting for this moment for a long time, and this effort is dedicated to them.

Being away from my family for many years, I learned to appreciate friendships much more. I am grateful to all my friends for their support, best wishes, and for their words of wisdom I needed to hear from time to time. I would like to extend my special thanks to Alexander Avramov for all the time he spent helping me with the statistical procedures and for always being there for me.

Chapter 1

INTRODUCTION

Conceptual Knowledge and Problem-Solving in Engineering Education

Engineering conceptual knowledge base is comprised of two major components: the fundamental physical principles and analytical procedures. Fundamental physical principles involve understanding how a system operates and why. Analytical procedures in engineering education are the mathematical equations that are essentially symbolic representations of the physical principles described above. The major purpose of these representations is to serve as a vehicle to solve problems in engineering. The combination of physical principles and analytical procedures creates the basis for engineering problem solving. However, for many engineering students even fairly successful application of analytical procedures to problem-solving doesn't necessarily translate into understanding of the physical phenomena that these procedures represent. Eventually, this incomplete conceptual understanding hinders the development of central engineering competencies described further (Streveler et al., 2004; Streveler et al., 2008).

Conceptual knowledge is "knowledge of categories and classifications and the relationships between and among them" (Anderson & Krathwohl, 2001, p. 48). It includes schemas, mental models, and theories that represent an individual's understanding of how a particular subject matter is structured and organized, as well as how the parts of a system are interrelated and function together (p. 48). One of the subtypes of conceptual knowledge that is especially important in engineering is knowledge of principles and generalizations. As Anderson & Krathwohl (2001) note, principles and generalizations require "knowledge of particular abstractions that summarize observations of phenomena" (p. 51). According to Streveler et al. (2008), conceptual knowledge is critical to the development of engineering competencies, as engineers rely heavily on conceptual knowledge during their engineering practice. One of the important aspects of engineering expertise includes intuitive expectations as to how systems act without having to refer to complex models or physical prototypes. This mode of operation is often called engineering judgment or heuristic thinking. To be able to engage in such thinking activities, an engineering professional has to have very strong conceptual foundations which primarily include understanding of basic quantities and understanding of relationships among the basic quantities (Streveler et al., 2008, p. 280). Conceptual knowledge, that is, knowledge of principles and generalizations in particular, has a great value in determining the most appropriate action that has to be taken (Anderson & Krathwohl, 2001, p. 51), which makes it especially important for problem-solving based on heuristic thinking.

Lovett (2002) defines problem solving as “the analysis and transformation of information toward a specific goal” (p. 317). This transformation requires the problem solver to take certain steps or follow certain procedures to achieve the goal. According to Mayer & Wittrock (2006), problem solving is defined by the following characteristics: (1) it is cognitive, because it occurs internally in the problem solver's cognitive system, and can only be observed indirectly from the person’s behavior; (2) it is a process, as it involves representing and manipulating knowledge in the problem solver's cognitive system; (3) this process is directed, which means that it is guided by the problem solver's goals; and, finally, (4) problem solving is personal, that is, it relies on the individual knowledge and skills of the problem solver (Mayer & Wittrock, 2006, p. 287). Engineering problem-solving for novice learners is predominantly well-defined. In a well-defined problem, “the given state, the goal state, and the procedures that should be used to reach the goal are clearly stated” (Mayer & Wittrock, 2006, p. 288). To summarize, problem solving is a complex process, which requires a particular set of skills from the learner, and these skills are both cognitive and procedural. These skills are not innate to the learner, and have to be acquired as a result of conceptual learning and procedural practice in a specific domain.

Worked Examples as a Learning Strategy for Engineering Problem-Solving

The conventional practice of teaching and learning problem solving in engineering education relies on worked examples presented during the lecture or in the textbook and practice problems learners go through in their homework. Worked examples usually model the solution process and consist of the problem statement, solution steps, and the final solution to the problem (Moreno, 2006). Worked examples help learners see similarities between the example and practice problems that need to be solved. Such ability to make connections based on similarities between several cases is called analogical mapping (Holyoak, 2005). The downside of this approach is that very often students are unable to make analogy-based connections between a worked example and practice problems that go beyond the surface features (e.g. symbolic representations), which can negatively affect transfer to similar problems as well as students' ability to solve more complex problems.

Initial studies conducted by Sweller and his colleagues (Sweller & Cooper, 1985; Cooper & Sweller, 1987; Tarmizi & Sweller, 1988; Ward & Sweller, 1990) demonstrated clear superiority of example-based learning over conventional problem-solving practices for schema development and rule automation. These studies also reported decrease in cognitive load due to elimination of means-ends analysis (Sweller et al., 1989). However, when the complexity of worked examples increased in the subsequent studies, due to either more complex principles involved, or use of instructional supports, or new computer technologies, the evidence of their effectiveness became less convincing (e.g. Conati & VanLehn, 2000; van Gog et al. 2006; Mahan, 2007).

The first main learning problem for novices with the way worked examples are typically used in engineering education, which create difficulties for students during problem-solving, is that novice learners have difficulty mapping the attributes of practice problems with those of worked examples, when worked examples are separated from practice problems in time and space (Atkinson et al., 2000). Making the connections between different features of the worked examples and practice problems more explicit may be one way to facilitate learning from worked examples for this type of learners.

Self-Explanation as a Self-Regulated Learning Strategy

Self-regulation refers to “self-generated thoughts, feelings, and actions that are planned and cyclically adapted to the attainment of personal goals (Zimmerman, 1999, p. 14). Self-explanation fits into this definition as it is, according to Roy & Chi (2005), “a domain-general constructive activity that engages students in active learning and ensures that learners attend to the material in a meaningful way while effectively monitoring their evolving understanding” (p. 272). Initial studies on effects of self-explanations for both worked examples and expository texts demonstrated positive effects on learners’ understanding of principles and procedures (Chi et al., 1989; Chi et al. 1994; Renkl, 1997), but further investigations brought inconsistent findings on self-explanation effects, especially when combined with other instructional supports (e.g. Schworm & Renkl, 2006; Gerjets et al., 2006; Berthold & Renkl, 2009). Some studies also revealed that use of self-explanations increased subjective cognitive load (Huang, 2007; Berthold & Renkl, 2009).

This is the second main learning problem for novices with the way worked examples are typically used in engineering education: without being prompted to make the connection, novices mostly attend to surface features in the examples, rather than deeper conceptual knowledge (Alexander, 2003). In engineering education surface features are usually represented by analogical procedures and conceptual knowledge is represented by physical principles (Streveler et al., 2004).

Multimedia Instruction as a Facilitation Tool for Multimedia Learning

According to Mayer (2005a), multimedia learning is "building mental representations from words and pictures" (p. 2), and multimedia instruction refers to designing multimedia presentations that promote building these mental representations (p. 3). In recent years, development of advanced computer technologies, such as advanced graphics and animations, allowed educators to take advantage of more complex visual ways of presenting material (p. 3). In science and engineering education the use of video and animations enabled educators to demonstrate physical principles and visualize fundamental molecular-level and atomic-level phenomena. At the same time, many undergraduate courses in engineering carry at least some part of their activities using a course management system, which serves as a learning environment in which these multimedia presentations can be delivered outside of classroom. In this study, current computer technologies are used to facilitate building mental representations of physical principles and analytical procedures and integrate them in the online learning environment.

Purpose of the Study

The purpose of this study was to investigate the effects of segmented multimedia-based worked examples and self-explanation prompts (SEP) on conceptual learning and problem-solving in an engineering course.

The first aspect of investigation explored two main effects of segmented worked examples and self-explanation prompts on acquisition of conceptual learning and problem-solving performance.

The second point of investigation explored possible interrelationships among acquisition of conceptual learning and problem-solving performance and type and quality of generated self-explanations.

Significance of the Study

This study builds on prior research on learning from various types of worked examples with and without additional instructional supports. From a theoretical stand-point, it extends understanding how worked examples and self-explanation prompts influence conceptual knowledge and problem-solving in the engineering domain. It also explored the extent to which explicit mapping of example attributes to the practice problem was necessary for successful learning from the worked example. This study also used new multimedia technologies that allowed rich representation of the example problem.

Another aspect of this study was further exploration of the role of individual self-explanation strategies in learning from worked examples in the domain that used complex physical and analytical principles.

There are several aspects of this study that have practical significance. First of all, this study was conducted in a real-classroom setting. This was one of the areas in example-based research that had not been sufficiently explored yet (Renkl, 2005). While it is possible to design a laboratory experiment that will create the perfect conditions for studying with worked examples, the real setting introduces a large number of variables the researchers, instructional designers, and course instructors need to account for in their practice. This study informs instructional designers and course instructors of the ways to support learning from worked examples in the classroom. Secondly, it demonstrates how multimedia learning principles (Mayer, 2005) can be applied to the design of example-based learning. Thirdly, it suggests ways to support the problem-solving process with metacognitive activities using content-specific self-explanation prompts in an online environment.

Research Questions and Hypotheses

Effect on Conceptual Knowledge

Question 1: Do different types of multimedia worked examples influence acquisition of conceptual knowledge?

Hypothesis 1: Different types of multimedia worked examples will significantly influence acquisition of conceptual knowledge.

Hypothesis 1-1: Learners who study with segmented multimedia worked examples enhanced with self-explanation prompts will demonstrate significantly higher conceptual knowledge score than learners who study with segmented multimedia worked examples alone.

Hypothesis 1-2: Learners who study with segmented multimedia worked examples will demonstrate significantly higher conceptual knowledge score than learners who study with non-segmented multimedia worked examples.

Effect on Problem Solving

Question 2: Do different types of multimedia worked examples influence problem solving performance?

Hypothesis 2: Different types of multimedia worked examples will significantly influence problem solving performance.

Hypothesis 2-1: Learners who study with segmented multimedia worked examples enhanced with self-explanation prompts will demonstrate significantly higher problem solving scores than learners who study with segmented multimedia worked examples alone.

Hypothesis 2-2: Learners who study with segmented multimedia worked examples will demonstrate significantly higher problem solving scores than learners who study with non-segmented multimedia worked examples.

Effect on Quality of Self-Explanations

Question 3: Is there an interrelationship between the type and quality of self-explanations, acquisition of conceptual knowledge, and problem solving performance?

Hypothesis 3: There is a significant interrelationship between the type and quality of self-explanations, acquisition of conceptual knowledge, and problem solving performance.

Definition of Terms

Worked example is a learning tool that is used to teach problem-solving skills in well-structured domains, such as physics or mathematics. Worked examples usually consist of modeling the process by presenting an example problem and demonstrating the solution steps, as well as the final answer to the problem (Moreno, 2006, p. 170).

Multimedia worked example in this study is a worked example that is developed using multimedia authoring tools to facilitate building mental representations of physical principles and analytical procedures (Mayer, 2005a) and is delivered online via a course management system. In addition to the traditional definition of a worked example (Moreno, 2006), the multimedia worked example in this study also contains instructional explanations, i.e. “pedagogical actions reflecting an attempt to give answers to questions that are implicitly or explicitly posed by learners or teachers” (Wittwer & Renkl, 2008, p. 50), which provide additional explaining to connect the principles and solution steps.

Segmented multimedia worked example in this study is a multimedia worked example that is integrated with the practice problem by breaking the continuous multimedia presentation into nine separate parts that show only the conceptual and procedural information related to each step in the practice problem to facilitate analogical reasoning (Holyoak, 2005).

Self-explanation prompt is a question whose purpose is to induce meaningful learning strategies, such as explicating the conditions and consequences of each procedural step in the example, as well as justifying the principles and definitions of concepts applied to them (Chi et al., 1989; Schworm & Renkl, 2007). In this study, self-explanation prompts elicit learners’ reasoning on how and why they applied certain principles in their solutions of the practice problems. Learners are asked to generate self-explanations explicitly, i.e. by typing them in the text box in the online environment (course management system).

Conceptual knowledge is “knowledge of categories and classifications and the relationships between and among them” (Anderson & Krathwohl, 2001, p. 48). It includes schemas, mental models, and theories that represent an individual’s understanding of how a particular subject matter is structured and organized, as well as how the parts of a system are interrelated and function together (p. 48).

Problem solving is “the analysis and transformation of information toward a specific goal” (Lovett, 2002, p. 317). This transformation requires the problem solver to take certain steps or follow certain procedures to achieve the goal. In this study problems are of the well-defined type. In a well-defined problem, “the given state, the goal state, and the procedures that should be used to reach the goal are clearly stated” (Mayer & Wittrock, 2006, p. 288).

Quality of self-explanations in this study is the extent to which individual learners engage in identifying and elaborating on the principles, provide the rationale of why the principle is used, and do not demonstrate systematic misunderstanding of underlying domain principles (Berthold & Renkl, 2009) in written self-explanations.

Chapter 2

LITERATURE REVIEW

Development of Expertise

Expertise is generally viewed as "consistently superior performance on a specified set of representative tasks for a domain" (Ericsson & Lehmann 1996, p. 277). Expertise in a well-structured domain, such as engineering, requires a well-developed set of problem-solving skills that can be readily applied to various situations. The correctness of application of these skills depends on solid understanding of the conceptual principles behind problem-solving procedures. Information processing plays the major role in acquisition of conceptual knowledge and problem-solving skills. This section reviews theoretical foundations of the cognitive processes that are involved in initial stages of expertise development.

Cognitive Models of Expertise

Based on Model of Domain Learning (MDL) introduced by Alexander (2003), there are three main stages of expertise development: acclimation, competence and proficiency, or expertise. In the acclimation stage novice learners have limited and fragmented knowledge; and they frequently use surface-level strategies; and rely on situational interest to maintain focus. In the competence stage, learners already possess a foundational body of domain knowledge, their knowledge is also more cohesive and principled in structure; learners apply a mix of surface-level and deep-processing strategies; and there is an increase in individuals' personal interest in the domain and less dependence on situational features of the environment. In the highest proficiency/expertise stage, knowledge base become both broad and deep, at that point experts are also contributing to building new knowledge in the domain; the level of strategy use among experts remains high and almost exclusively of a deep-processing kind; and the individual interest of experts is very high, while reliance on situational interest levels off.

Based on the model mentioned above, there are several directions in which changes occur as a learner moves from novice to expert. First of all, the knowledge base becomes significantly broader and deeper. Then, level of processing changes from surface to deep, or, using different terminology, it changes from lower-order to higher order thinking skills. Motivation changes from extrinsic to intrinsic. There are also important changes that happen on the metacogntive level: ability for self-assessment increases as the level of expertise increases (van Gog et al., 2005).

A four-stage model of expertise, developed by Anderson (1995), is consistent with schema theory and is based on the Adaptive Character of Thought (ACT-R) framework (Anderson, 1995; Anderson et al., 1997). According to the framework, at the first stage of a skill acquisition, learners solve problems by analogy, which is relating the known examples to the problem that they need to solve. At the second stage, learners develop declarative encoding of the skill, or a schema, and typically rehearse the set of facts as they perform the skill. At the third stage, the schemas become more proceduralized, as connections among various elements of the skill are strengthened through practice. At the fourth stage, on the one hand, schemas become automatic, and on the other hand, the sufficient level of expertise and experience allows for operating from a large pool of various types of examples using a higher level of analogical reasoning.

Implications for the Current Study

To reach the level of expertise, novice learners need to build their knowledge base, develop their deep-processing strategies and learn how to reflect on their progress. While acquiring expertise takes years of practice, the instruction that supports acquisition of principles and relations among principles, and also provides self-reflection opportunities, should be especially helpful to novice engineering students in moving through the acclimation stage and reaching the level of competence. The four-stage model of expertise helps proceduralize this process. In accordance with this model, this study will employ segmented worked examples to facilitate analogical reasoning and development of schemas to build foundations of engineering expertise.

Analogical Reasoning Theory

Analogical reasoning is a process that allows connecting a new situation with a previously experienced situation by establishing the similarity between the two through a set of inferences. According to Holyoak (2005), analogical reasoning consists of a “target” situation that serves as a retrieval cue for a potentially useful “source.” To connect the target with the source and determine their analogous nature, it is necessary to establish a “mapping”, which is a set of correspondences between the two cases. Once the similarities are established, the transfer from source to target occurs. Based on mapping, it is possible to draw new inferences about the target and elaborate its representation. As a result of this process, such mapping could lead to development of a more abstract schema which in its turn could be applied to different situations, thus promoting analogical transfer. Mapping can be guided depending on the goals of the instruction, which means that we can guide learners to attend to certain features and relations of the source case, rather than the others.

Analogical mapping depends on manipulating structured representations of knowledge (Holyoak, 2005). Structured representations explicitly encode information about the scope of representational elements (Markman, 2002). That means that such representations not only contain information regarding the features of the representation, but also specify the relationships between their elements. Correct analogical inference happens when the target is mapped to the source on the basis of relations between the elements rather than on featural similarity (Holyoak, 2005).

Implications for the Current Study

Analogical reasoning is one of the most crucial underlying mechanisms that allow worked examples to be effective in the learning process. From ACT-R framework, described earlier, we can see that novice learners need to rely on analogy to solve problems. To help novices learn from worked examples more effectively, this study will use segmented worked examples, where elements of the worked example will be linked to the corresponding elements of the practice problem, to facilitate analogical reasoning by providing more opportunities to establish the correct analogical inferences and emphasizing the conceptual relations between the worked example, which serves as a source for mapping, and the new problem that has to be solved, which in this case is the mapping target.

Cognitive Load Theory

Cognitive load theory is based on the assumption that human working memory has limited capacity: according to Miller (1956), it can only hold around 7 chunks of information simultaneously. At the same time, long-term memory capacity is virtually unlimited. The capacity of working memory can be increased if information is organized into schemas. Schemas serve both for storing learned information in long-term memory and reducing the cognitive demands on working memory (Kalyuga, Chandler, & Sweller, 1998). According to Sweller et al. (1998), schema construction and rule automation are two main principles that guide knowledge acquisition for novice learners. Intellectual skill consists of storing a high amount of schemas of high levels of automaticity in the long term memory (p. 258). These schemas can be easily retrieved from the long term memory to ensure fluent performance of the skill.

Before schemas reach long term memory, they need to be processed in working memory. Cognitive load theory is mostly concerned with the ease of processing information in working memory (Sweller et al., 1998). With respect to the processing demands, instructional materials carry intrinsic cognitive load, which is imposed by the nature of the material, or extraneous cognitive load, which is determined by the manner in which material is presented (p. 259). Extraneous cognitive load is usually associated with unnecessary cognitive demands that are not directed at schema acquisition (Sweller et al. 1998; Moreno, 2006). That is the reason why extraneous cognitive load should be differentiated from germane cognitive load, which is evoked by instructional activities directly related to schema construction and automation (Sweller et al. 1998; Moreno, 2006).

Empirical research on cognitive load theory allowed for discerning various effects that could be potentially harmful for learning, such as split-attention effect (Tarmizi & Sweller, 1988;

Ward & Sweller, 1990), redundancy effect (Kalyuga et al, 2000; Sweller, 2005), and expertise reversal effect (Kalyuga et al, 2001).

Paas & van Merrienboer (1994) developed a 9-point mental effort scale, which was designed to measure "the cognitive capacity that is actually allocated to accommodate the demands imposed by the task" (Paas et al., 2003, p. 64), i.e. the actual cognitive load. This scale does not differentiate among different types of cognitive load. DeLeeuw & Mayer (2008) determined that extraneous processing could be measured by response time, intrinsic processing influenced effort ratings, and difficulty ratings were indicative of germane processing.

Implications for the Current Study

While it is not completely clear how to measure different aspects of cognitive load, managing demands on working memory is very important for learning from worked examples. Worked examples are believed to decrease cognitive load, but only if all possible negative effects, such a split-attention or redundancy effects are taken into consideration. Use of worked examples in this study as opposed to conventional problem-solving strategies, should help decrease extraneous cognitive processing. This effect should be increased by use of multimedia presentations, which help avoid split-attention effect by combining verbal and non-verbal representations. At the same time, self-explanations are believed to increase cognitive load. To avoid redundancy effect, they will be used once per problem.

Multimedia Learning

In recent years, employing different media of various levels of complexity and interactivity became a widely-used strategy in instructional design. The majority of the current research reviewed in this paper was done using computers either to deliver the intervention (e.g. Atkinson et al., 2003; Renkl et al. 2004; Reisslein et al., 2006) or to support a complex interactive system (e.g. Conati & VanLehn, 2000; Aleven & Koedinger, 2002; Darabi et al., 2007). Delivering the intervention in an electronic format has both advantages and caveats, which will be discussed in this section.

Dual Coding Theory and Multiple External Representations

Paivio (1991) lays out the main assumptions of dual-coding theory. According to Paivio, verbal and non-verbal systems symbolically represent the structural and functional properties of language and the non-linguistic world. At the same time, both classes of events can be represented in different modalities: visual (printed words vs. visual objects), auditory (spoken words vs. environmental sounds), and several types of sensorimotor modalities. The main emphasis of the theory is on the importance of verbal vs. non-verbal symbolic contrast, which means that the same object can be characterized using words (printed or spoken) or images/sounds/manipulation of objects.

Since the information from sensory input can be received and encoded from both verbal and non-verbal channels using different modalities, the same concept or event can be represented in memory using multiple representation modes. The same concept can be described using printed words or an image, or a video, or a graph, or an equation. These external representations can either complement, i.e. support each other, because they contain different pieces of information; constrain, i.e. one representation can constrain the interpretation of the other; or construct, i.e. build on each other to support construction of deeper understanding (Ainsworth, 2006). Very often one representation of a concept is not enough, and it is necessary to help students to understand and interpret information from different external representations, as well as to integrate these interpretations to develop deeper and more complete understanding of the concept.

Principles of Multimedia Learning

Multimedia learning presents educators and researchers with interesting and promising capabilities, because it allows implementation of more complex instructional solutions that were either not possible or cumbersome using more traditional means of instruction, such as textbooks and face-to-face classroom learning. Computers also make testing easier to implement.

Existing research in cognitive theory provides important foundations for effective multimedia learning design. Mayer (2005b, 2005c, 2005d) reviewed the current theories and empirical evidence in support of three types of principles that could potentially influence the effectiveness of learning with multimedia.

The first set of principles is aimed at managing essential processing, and includes segmenting, pre-training, and modality principles (Mayer, 2005b, p. 169). Usually multimedia environments are designed to accommodate complex content. It also provides a means to integrate multiple representations more effectively (Moreno & Mayer, 2007). However, these overall positive characteristics of multimedia learning usually lead to presenting a large amount of essential material at a rapid rate, which may result in essential overload (p. 170). Since human working memory capacity is limited (Baddeley, 2003), overloading it even with essential information is not productive. To effectively manage essential processing, multimedia materials should: (1) be segmented, i.e. break the presentation in smaller pieces that allow learners control over the pace of instruction (Mayer, 2005b, pp. 171-173); (2) use pre-training when possible; and (3) be multimodal, i.e. combine both verbal and non-verbal representations of knowledge (Low & Sweller, 2005; Moreno & Mayer, 2007).

The second set of principles deals with reducing extraneous processing, and refers to coherence, signaling, redundancy, and contiguity principles (Mayer, 2005c, p. 183). Multimedia environments are typically richer than paper-based instructional materials, but this richness can also translate into presence of visual and audio cues that are not necessary for learning. These non-relevant cues can potentially result in extraneous overload (p. 184). To reduce extraneous processing, multimedia materials should be (1) coherent, i.e. exclude extraneous material; (2) use signaling, i.e. include cues that highlight the organization of essential information; (3) non-redundant, i.e. use graphic and narration, rather than graphic, narration, and on-screen text; and (4) follow spatial and temporal contiguity principle, i.e. present corresponding words and images near each other and at the same time (pp. 184-185).

The third set of principles is based on social cues, such as personalization, voice and image. Current multimedia capabilities allow using social cues of various level of fidelity. However, according to empirical data, people learn better when the words in the multimedia presentation are in conversational rather than formal style, when they are spoken in a standard-accented human voice, and presentation of the speaker's image on the screen does not necessarily influence learning (Mayer, 2005d, p. 201).

Implications for the Current Study

For the instruction to be more effective, this study will use several modes of representation of information, such as text and video, as well as several modalities, such as written text and audio. The multimedia worked examples will avoid irrelevant and redundant material, use labeling and highlighting as a signaling strategy, and present corresponding words and images near each other and at the same time to reduce extraneous processing. The information will also be presented using the course instructor's voice and conversational format.

Worked Examples

Definition and Principles

Worked examples are instructional tools that are used to teach problem-solving skills in well-structured domains, such as physics or mathematics. Worked examples usually consist of modeling the process by presenting an example problem and demonstrating the solution steps, as well as the final answer to the problem (Moreno, 2006, p. 170). In an example-based model, past examples are retrieved from memory based on their similarity to the current problem-solving situation. Then the information stored in the example is used to determine how to approach the new problem (Lovett, 2002).

Use of worked examples in problem-solving is consistent with the four-stage model of expertise that is based on the Adaptive Character of Thought (ACT-R) framework (Anderson, 1995; Anderson et al., 1997). According to this model, worked examples are most useful at the first two stages of expertise development, when learners need to develop initial schemas to be able to solve similar problems. As the learners' level of expertise increases, they will rely less on worked examples (Kalyuga et al. 2001) and more on already developed schemas (Sweller et al., 1998).

Relevant Studies on Worked Examples

Research on worked examples started in 1980s with a set of experiments conducted by Sweller in association with various researchers (Sweller & Cooper, 1985; Cooper & Sweller, 1987; Tarmizi & Sweller, 1988; Ward & Sweller, 1990). These studies demonstrated superiority of worked examples over conventional problem solving practice for similar, or near transfer, problems, and also provided empirical foundations for schema acquisition and rule automation, as the initial stages of development of expertise. According to Sweller & Cooper (1985), use of worked examples eliminates the need for learners to engage in means-ends analysis, an unstructured activity in which a person "attempts to reduce differences between each problem state encountered and the goal state" (pp. 61-62). This effect is achieved due to redirecting the learner's attention from the goal state towards structured problem-state configurations, which results in decreased amount of time necessary for acquisition of problem-solving schemas (Sweller & Cooper, 1985, p. 86). These findings were confirmed and extended to rule automation and worked example format effect in the subsequent studies (Cooper & Sweller, 1987; Tarmizi & Sweller, 1988; Ward & Sweller, 1990). It was established that initial schema development preceded rule automation, and worked examples facilitated both, unless they were presented in a split-attention format. These findings were attributed to the decrease in cognitive load, which was measured in time to solution for both schema acquisition and retrieval.

Research on effects of worked examples on analogical reasoning and working memory load was continued by Robins & Mayer (1993), who found that learners exposed to worked examples that were schematically structured and presented in a low-load format performed better on the near transfer test and worse on the memory test. These findings were consistent with the active learning theory, which is "based on the cognitive view that learning occurs when the learner is mentally active during learning and the teacher's role is to assist the learner in constructing knowledge, such as relational schemas" (p. 530). Also, consistent with this theory, schematic/low-load group didn't demonstrate higher transfer on dissimilar problems, i.e. far transfer.

To understand the underlying mechanisms that influence near and far transfer, it is first necessary to review their definitions. Near transfer problems are problems with the same structure or solution rationale but different surface features such as objects and numbers (Atkinson et al. 2003). Near transfer can be achieved by consistent mapping of surface features between the example and the problem (Holyoak, 2005), and initial schema development and rule automation (Anderson et al., 1997; Sweller et al., 1998). Far transfer problems, on the opposite, are problems with different structure that required the generation of a modified solution procedure (Atkinson et al. 2003). Far transfer is also sometimes called high road transfer. The main distinction of the high road transfer is "the mindful generation of an abstraction during learning and its later application to a new problem or situation from which basic elements are similarly abstracted" (Salomon & Perkins, 1989, p. 127). According to Salomon & Perkins (1989), mindful abstraction is "a deliberate, usually metacognitively guided and effortful, decontextualization of a principle, main idea, strategy or procedure" (p. 126). This definition suggests that far transfer doesn't happen unless (1) there is deep conceptual understanding of principles and procedures; and (2) such understanding is induced and supported by active metacognitive processing.

Since the classical worked-example research demonstrated the need to find solutions to support conceptual learning and far transfer, more recent studies attempted to resolve this issue using a variety of strategies that could potentially facilitate acquisition of deeper underlying structures from the worked examples.

Van Gog et al. (2006) assessed the effects of adding process information to worked examples that were designed to teach troubleshooting of electrical circuits. It was hypothesized that studying worked examples would lead to better transfer performance than solving conventional problems, with less investment of time and mental effort during training and test. The results showed that while worked examples were clearly superior to conventional problem-solving, adding process information resulted in increased investment of effort during training, but not in higher transfer performance when they were combined with worked examples. A similar study was conducted by Darabi et al. (2007), who explored the effects of process-oriented and product-oriented worked examples on performance, mental effort, and learners' involvement, where the latter was operationalized as a function of the first two. The researchers used a computer-based environment simulating troubleshooting of a chemical plant. The process-oriented group received information about underlying principles involved in troubleshooting, while the product-oriented group received information about troubleshooting procedures. The results revealed that learners' involvement for worked example groups was significantly higher than in control group, but no differences in learners' involvement score between two experimental conditions were found. At the same time, product group outperformed the process group in learning.

In a similar vein, Mahan (2007) studied the effects of principle-based and procedure-based worked examples on undergraduate students' transfer of principles. In that study, the procedure-based group was asked to solve the problem following the procedurally-structured worked example, and the principle-based group was asked to explain the solutions that were already worked out, based on the model in the worked example. The assumption that principle-based worked examples would be more effective for transfer of principles was not confirmed.

Another strategy employed by researchers was the use of instructional elaborations to support conceptual and procedural knowledge acquisition. Catrambone & Yuasa (2006) reported the results of the experiment conducted in the domain of computer programming, more specifically, writing database queries. The study was measuring the effects of active learning and types of instructional elaborations. In the active learning conditions, after reviewing a worked example, students were asked to fill-in the blanks in an exercise. In elaboration conditions, after reviewing a worked example, students were given an instructional elaboration. Although the authors hypothesized that excessive instructional information for more knowledgeable learners could hurt performance, because of the redundancy effect, the results indicate that conditions that had a structured exercise and elaboration or elaboration-only improved procedural performance the most. At the same time, active learning conditions required more training time. Some of the other studies involving different kinds of instructional explanations and elaborations will be discussed in the next section of this review.

Another aspect of worked-example research concentrates on discerning the effects of various combinations of worked examples in relation to practice problems. Reislein et al. (2006) used different types of worked examples, i.e. example-problem pairs, problem-example pairs, and fading in a series of worked examples, also controlling for prior knowledge in the domain of series and parallel electrical circuit analysis. No differences between the instructional procedures were observed, but low prior knowledge learners benefited most from traditional example-problem pairs while high prior knowledge participants learned most from problem-example pairs. This finding was explained by possible expertise-reversal effect, when more knowledgeable learners would find the examples redundant.

In addition to individual studies, Atkinson et al. (2000) reviewed a number of empirical studies on worked examples and came to the conclusion that worked examples were more effective when they: (1) were in proximity to matched problems; (2) were integrated to avoid split-attention effect, e.g. video and text are presented in same window; (3) used multiple modalities, e.g. audio and video, or audio and text; and (4) used a clear subgoal structure, meaning that each step had to be labeled (Atkinson et al. 2000).

Implications for the Current Study

According to the existing theoretical and empirical literature, worked examples are superior to conventional problem solving in learning problem solving procedures, especially for novice learners, as they facilitate initial schema development through analogical mapping of the worked example attributes. Worked examples also reduce cognitive load, as they decrease extraneous cognitive processing by eliminating means-ends analysis (Sweller & Cooper, 1985).

Also, using worked example – problem pairs and series of worked examples is more effective than using one worked example followed by a set of practice problems (Renkl, 2005). However, at the same time, there is conflicting evidence whether worked example – problem pairs can provide enough support for developing of deeper conceptual understanding, which is necessary to transfer problem solving schemas to novel problems (Sweller, 1985; Robins & Mayer, 1993; Grosse & Renkl, 2006; van Gog et al. 2006; Berthold & Renkl, 2009). This study will pair the worked example with the practice problems, and will support deeper conceptual understanding with self-explanation prompts described in the next section.

Worked Examples Enhanced with Self-Explanations

Self-explaining as a Self-Regulated Learning Strategy

The term *self-explanation* was first coined by Chi et al. (1989) in a study that assessed students' think-aloud protocols while they were studying worked examples in mechanics. Self-explaining in example-learning is a mechanism of study that allows students to infer and explicate the conditions and consequences of each procedural step in the example, as well as to apply the principles and definitions of concepts to justify them (Chi et al., 1989).

Self-explaining allows learners to monitor their understanding, and serves as a powerful self-regulated learning strategy. According to Chi et al., (1989), self-explanations can help students' understanding by making them think about at least one of the following: (1) the conditions of application of the actions; (2) the consequences of actions; (3) the relationship of actions to goals; or (4) the relationship of goals and actions to natural laws and other principles.

Some learners can generate self-explanations spontaneously as they study a text or an example. But as Renkl (1997) found, the majority of learners would not engage in self-explaining unless they were prompted to do so. A prompt is a question or elicitation which purpose is to induce meaningful learning activities (Schworm & Renkl, 2007). Usually these activities involve a set of learning strategies that students are capable of but don't typically engage in on their own (Renkl, 1997; Schworm & Renkl, 2007). Prompting helps students initiate active processing of information and directs their attention to the aspects of the instruction that are the most critical for learning.

According to Roy & Chi (2005), individual differences play an important role not only in learners' readiness to engage in self-explaining, but also in the quality of generated self-explanations, which can influence learning outcomes.

Relevant Studies on Self-Explanation Effects

In their study, Chi et al. (1989) noticed that "the discrepancy between those studies which show that students perform better from text materials that contain examples, and those studies which show that students often fail to generalize from examples, may be caused by the degree to which students understand the examples provided" (p. 148), and hypothesized that students learned and understood the examples through explaining it to themselves while they were studying them. The results of the transfer tests and think-aloud protocol analysis demonstrated that students who scored the highest on the tests were also the ones who had a tendency to study examples by explaining them and providing justifications for their actions (p. 175). In another study conducted several years later, Chi et al. (1994) found similar effects for students studying from expository texts. In that study, the researchers were comparing self-explanation prompts against the control group. The learning gains were significantly higher for the prompted group with high explainers also scoring higher than low explainers on tasks that required deeper conceptual understanding: knowledge integration and generation.

The implications of these studies were further investigated by Renkl (1997), who also analyzed students' think-aloud protocols, which were generated when students were engaged in example learning in the domain of probability calculations. He found that learning gains could be to a large extent predicted by qualitative differences in self-explanations. While Renkl (1997) admitted that individual self-explaining characteristics were multidimensional, he made two major conclusions: (1) successful learners tended to use more principle-based explanations, more explication of operator-goal combinations, and more anticipative reasoning; and (2) all observed effective learners could be labeled either anticipative reasoners or principle-based explorers.

Ainsworth & Loizou (2003) investigated whether the representation format, i.e. text or diagram, made a difference in the amount of self-explanations. The results revealed that: (1) diagram condition performed significantly better on post-tests than text condition; (2) diagrams students also generated significantly more self-explanations that text students; and (3) number of self-explanations was significantly negatively correlated with time spent on learning and number of words (p. 676). In their analysis of think-aloud protocols, Ainsworth & Loizou (2003) used the categories developed by Renkl (1997).

These findings are consistent with research done by Atkinson et al. (2003). Atkinson et al. (2003) conducted two experiments, where they were comparing the effects of different combinations of backward fading and self-explanation prompts on near and far transfer from worked examples. They found that the use of self-explanation prompts in combination with a backward fading example positively affects learning. As a self-explanation prompt, the learners were asked to select an underlying principle for each step of the solution process. This research was expanded by Renkl et al. (2004) who investigated the differences between backward fading, forward fading, and problem-example pairs. No significant differences between types of fading were found, but in the second experiment backward fading was found more effective than problem-example pairs. At the same time, neither strategy was superior with regards to the quality of self-explanations.

As far as the types of self-monitoring prompts, Berthold et al. (2004) examined the effectiveness of different types of prompts for writing lesson follow-up learning protocols. They found that, out of cognitive, metacognitive, mixed prompts or no prompts, only cognitive and mixed prompts were effective with regard to the learning outcomes on an immediate comprehension test and a 7-day-delayed test.

Sometimes, positive effects of self-explanation prompts are attributed to increased cognitive load (Moreno, 2006). The nature of this cognitive load is assumed to be germane, i.e. conducive to learning. A number of researchers tried to shed more light on the issue. Gerjets et al. (2006) attempted to reduce intrinsic cognitive load by using modular worked examples, which were examples segmented into meaningful structural elements, as opposed to conventional molar examples, which were more cognitively demanding. At the same time, the authors tried to increase germane cognitive load by adding either instructional elaborations or a combination of elaborations and self-explanations. The results showed that modular format was far superior to the molar one; and it also reduced cognitive load. Neither instructional elaborations nor self-explanations significantly improved performance, and prompting for self-explanations even impaired learning with modular examples. The authors suggested that prompting might have forced learners to process the information that was already understood, thus leading to redundancy effect described in cognitive load literature.

Große & Renkl (2006) investigated different approaches to teaching combinatorics and probability using multiple-solution methods and different types of explanations. After the first experiment, it was concluded that multiple solutions fostered learning. However, these findings could not be replicated in the second experiment. Neither intervention revealed positive effects for instructional support. Multiple solutions method reduced some of the aspects of spontaneous self-explanation activities during think-aloud sessions. The authors attributed lack of positive outcomes to high intrinsic cognitive load of the subject matter and possible extraneous cognitive load induced by think-aloud procedures.

Hilbert et al. (2008) investigated the effectiveness of self-explanation prompts and completion of gaps when studying from heuristic worked examples. As a result of the study, students learning with heuristic examples outperformed their control group counterparts on both proving skills and conceptual knowledge Use of self-explanation prompts only fostered conceptual knowledge as well as skills, but completion of gaps in combination with self-explanation prompts significantly impaired learning. The authors hypothesized that “the demand to react to the prompts and gaps induced an excessive demand of essential processing that resulted in cognitive overload” (p. 63).

One of the most recent studies conducted by Berthold & Renkl (2009) investigated whether use of multiple representations was enough to support conceptual understanding in probability or other instructional supports, such as color-coded flashing learning aids and self-explanation prompts, were needed to fully use the potential of these representations. Some of the important findings were: (1) conceptual knowledge was fostered by scaffolding self-explanation prompts, but these prompts had a negative effect on procedural knowledge irrespective of the type of provided representation; (2) self-explanation prompts frequently led to confusion of principles; (3) both rationale-based self-explanations and principle-based self-explanations had positive effect on conceptual knowledge; and (4) learning without a relating aid increased subjective cognitive load, but scaffolding self-explanation prompts also increased load for some types of representation.

Some possible explanations of the nature of cognitive load evoked by self-explanations could probably be inferred from recent studies by Schworm & Renkl (2006, 2007). Schworm & Renkl (2006) investigated the effects of instructional explanations and self-explanations when studying how to develop effective worked examples by using solved examples. Schworm & Renkl (2007) explored the use of video-based worked examples in an ill-structured domain of argumentation skills. The study employed different combinations of learning-domain and exemplifying-domain prompts. Video-based worked examples consisted of dialogues that contained both learning domain and exemplifying-domain content, i.e. argumentative structure and contents of the dialogues respectively. Both studies reported positive effects of using self-explanation prompts. The first study also found that instructional explanations reduced the effects of self-explanations. But the most unexpected finding from both studies was that though self-explanations had objectively measured positive learning outcomes, students did not perceive them as being particularly helpful. Students who don't perceive self-explanation activities as useful will be more likely to treat them as extraneous processing. The question remains open whether objective benefits of self-explanations can outweigh the increased cognitive load in the minds of individual learners.

Another end of research is the use of cognitive tutors that employ self-explanation generation strategies. Studies, such as Conati & VanLehn (2000) and Aleven & Koedinger (2002), used different variations of computer-based environments that allowed adaptive self-explanation prompts. These studies did not produce conclusive results regarding the effectiveness of adaptive self-explanation prompting. Aleven & Koedinger (2002) reported positive effects of such programs, but Conati & VanLehn (2000) found evidence of self-explanation effect only based on prior knowledge, and not on instructional strategy. A more recent empirical research review by Koedinger & Aleven (2007) suggests that students using cognitive tutors rarely engage in self-explanations unless they are provided with feedback on them.

Implications for the Current Study

Self-explanation prompts seem to be a fairly effective strategy, especially when they incorporate a cognitive component, rather than a purely metacognitive one. At the same time, there is conflicting evidence regarding the effectiveness of self-explanation prompts depending on the content, type of problems, structure of worked examples, and presence of other instructional supports. Self-explanations seem to be necessary to promote conceptual understanding, but sometimes they can be perceived as redundant and not helpful by the students. To address these concerns, this study will use self-explanation prompts to elicit deeper understanding of the underlying domain principles, but at the same time will limit the amount of self-explanations to a necessary minimum to avoid redundancy effect.

Summary

Based on the literature above, problem-solving from worked examples was predicted to be more effective when the worked example is either integrated or in close proximity to the practice problems. At the same time worked example-based instruction benefits from using several modes of representation and using more than one example problem.

According to analogical reasoning theory, linking individual features of the practice problem, such as solution steps, to the corresponding features in the worked examples was also predicted to help learners establish similarities between the practice problem and the worked example, which should facilitate analogical transfer and develop the schemas that can be used to solve similar problems in the future.

At the same time, using a worked example alone may not be enough to help learners establish connections among procedural steps in the problem and conceptual understanding of the underlying engineering principles. The literature suggests that carefully designed self-monitoring prompts can be helpful in conceptual learning, especially if they include a cognitive component as well as a meta-cognitive one.

The instructional design framework for this study, shown in Figure **2.1**, integrates a worked example into the homework problem, thus creating an example-problem pairing, which should help novice learners make deeper connections between different attributes of both cases. At the same time, the self-monitoring component of this framework should help the novice learners attend to the relations among concepts rather than to the surface features of the example.

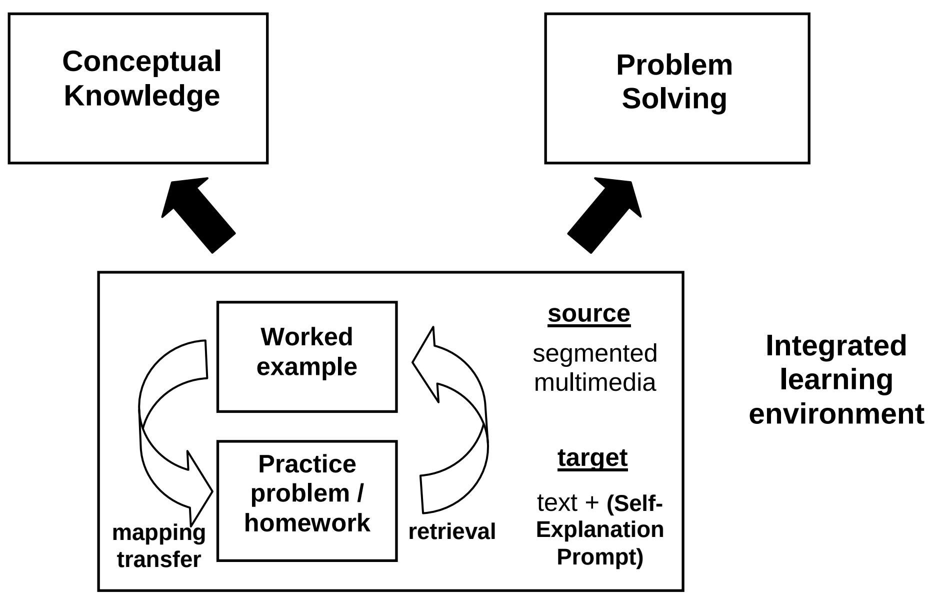


Figure 2.1. Design Framework

In the implemented design, a worked example was represented by an instructor-recorded multimedia presentation, which included video, and instructor voice and notes recorded using a multimedia authoring tool called Camtasia and a tablet-PC (a laptop computer that has advance input capabilities). This multimedia presentation was segmented into steps based on the problem-solving process the students had to engage in while solving homework and exam problems. These worked example chunks were called segmented multimedia in the study. This Integrated Learning Environment was designed to address the first problem, stated in the introduction, which was the difficulty of analogical transfer when the worked example was disconnected from the practice problem. The environment was built around the concept of analytical reasoning, where students needed to be provided opportunities for analogical mapping and transfer of the attributes from the source (the worked example) to the target (the homework problem). After the transfer had occurred, the attributes of the “mapped” source could later be retrieved for solution of similar problems. While such careful assisted attribute-to-attribute mapping is necessary for initial learning from a worked example, as the level of competence increases, it was expected that students should be able to engage in the mapping process with less and less assistance. To help learners gradually move to independent construction of inferences, the same worked example was paired with three practice problems of increased difficulty.

As another point of reference, the framework utilizes the concept of dual coding by providing the segmented worked example in the form of a multimedia presentation, which includes predominantly audio, video, and pictures. The homework problems, in which the segmented worked examples were embedded, were mostly textual. Multiple external representations were provided within the multimedia presentation through the use of graphs, equations, and real-world examples of the same concept.

To address the second concern, which was the novices’ lack of ability to attend to the deep conceptual structures while solving problems, self-explanation prompts were used as a self-monitoring strategy. Through open-ended questions, students were prompted to explain which engineering concepts they would apply while solving for a particular step.

Chapter 3

METHODOLOGY

Introduction

This study was designed to be conducted in real classrooms, and was conducted in two stages: the pilot and the main study. The purpose of the pilot was to refine the treatments and test the logistical procedures. It also attempted to answer the question whether segmented multimedia worked examples and self-explanation prompts influenced performance on problem-solving tests. The purpose of the main study was to explore the effects of segmented worked examples and self-explanation prompts on acquisition of conceptual learning and problem-solving performance and to explore the interrelationship among acquisition of conceptual learning and problem-solving performance and type and quality of generated self-explanations.

Pilot Study

The purpose of the pilot, which was conducted in Spring 2009, was to refine the treatments and test the logistical procedures. The research question asked in the pilot study was: do segmented multimedia worked examples and self-explanation prompts influence performance on problem-solving tests?

The participants of the pilot study were 67 engineering students enrolled in a 300-level Civil Engineering course. For all participants this was a required course for their major. While typically students at this course level are only beginning to specialize in civil engineering, and are assumed not to have substantial knowledge of engineering principles, it was expected that those who enrolled in the course would have solid knowledge of algebra and physics. Due to the specifics of engineering student population, and well as the university demographics, the majority of the participants were white male students around ages 19-22.

Materials and Treatments

The treatments were delivered as a part of the homework assignments for a 300 level Civil Engineering Course. Each problem-set consisted of a multimedia worked example on a fluids mechanics topic and 3 practice problems. The worked example was the instructor's explanation of the topic, which included general information about the physical phenomenon, analytical information, i.e. explanation of equations, and solution of an example problem divided into 4 steps.

The practice problems had the same structure as the example problem and consisted of 4-5 steps. Each step had a procedural label, e.g. 'assess the continuity equation', or 'calculate velocity'. Before starting the solution process, participants could view the problem statement with the link to the complete multimedia-based worked example, described above. The practice problems for the most part required students to manipulate surface features of the example problem, with slight deviations from the pattern for one of the steps in the last problem of the set. As participants went through the problem-set, they had to answer the test questions, which asked them to either calculate a certain parameter of the problem or compare two parameters.
The general set-up described in the previous paragraph was the basis for all four conditions.

In the Segmented Multimedia Worked Example condition, the multimedia worked example was segmented into 7 chunks. The first three segments presented the conceptual information regarding the problem, and the last four segments presented the four steps that were necessary to take to solve the problem. The segments then were linked to the steps of the practice problems.

In the Self-explanation Prompts condition, participants were prompted to explain their reasoning for each particular step. Self-explanation prompts presented an open-ended question which was located on the same screen as the calculation question(s) right before the calculation question. The wording of each SEP was tied to the conceptual component of the step, and was generally formulated as: “How would what you know about <a general principle involved in the step> help you solve for this step?” Participants were asked to type their self-explanations in a text box.

In the Segmented Multimedia Worked Example enhanced with Self-explanation Prompts condition, both segmented multimedia treatment set-up and self-explanation prompts treatment set-up were used. Figure **3.1** demonstrates an example of one of the steps in the Segmented Multimedia Worked Example enhanced with Self-Explanation Prompts condition.

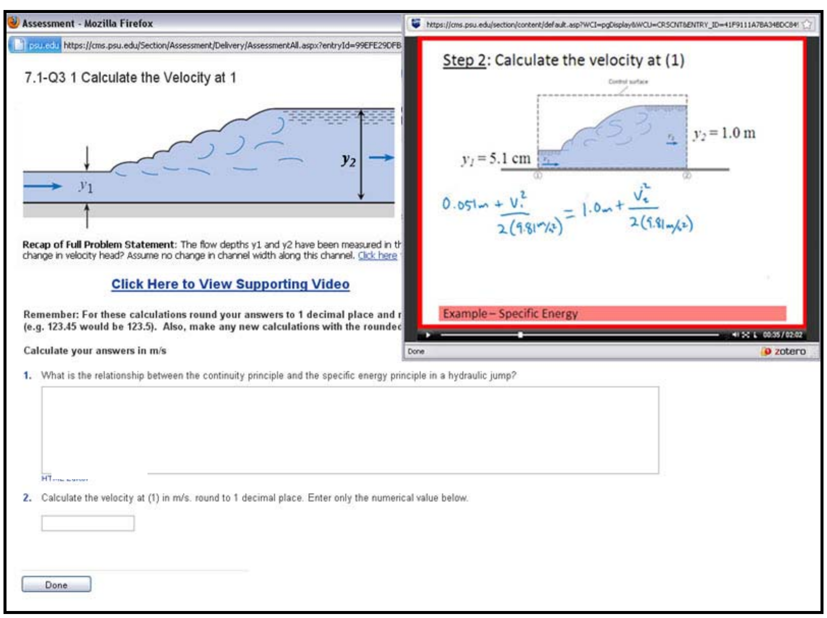


Figure 3.1. Pilot Study Segmented Multimedia Worked Example with SEP Screenshot

Measurement Instruments

Problem Solving:

Problem solving performance of students for the Specific Energy problem-set was measured with two types of problem-solving tests. The first test, called in-treatment problem-solving test, was designed as a part of the treatment materials and asked students to provide step-by-step solutions to 3 problems. The test consisted of 21 items generated from these three problems with the maximum score of 21. Problem 1 included 4 steps - 7 test items; problem 2 included 5 steps - 6 test items; problem 3 included 5 steps - 8 test items. The order of the questions in the homework was consistent with the step-by-step solutions of the problems. Participants typed in the answers to the test questions as they proceeded through the homework assignment in the course management system. The test administration on the learning management system was set for it to be completed in one sitting. Based on the results of the pilot study, Cronbach's Alpha reliability for the responses to this test was .79.

The second test called the delayed problem-solving test, consisted of 1 free scored problem with a maximum 4 points. Students were asked to write out a complete solution for a problem, which tests the same concepts as the homework problems, but was different enough to require a meaningful application of the principles, rather than simply following the procedures. This test was hand scored by a teaching assistant, who was a graduate student in engineering. The test consisted of 1 item, and the maximum score was 4. The results of this test were used only to inform the design, with understanding that this test would be revised before the main study to eliminate any reliability concerns, as it was impossible to determine internal consistency of the test that has only one item.

Design Survey

An end of semester survey was used to collect students' opinions about the course activities, including the study design to make changes before the next iteration. The survey included 14 questions in the following six categories: 1) attitude toward the overall design of the homework assignments; 2) frequency and reasons for viewing segmented video; 3) students' self-reported interactions with the treatments; 4) perceived effects of the treatments on students' learning; 5) perceived effects of the treatments on other aspects of the course; and 6) suggestions for improvement. Eight questions were Likert-type scale, 2 questions were checkboxes (check all that applies), and 4 questions were open-ended.

Procedures

The following procedures were completed during the pilot process:

- Two problem-sets were developed, each consisting of a multimedia-based worked example and 3 homework problems
- Participants were randomly assigned to one of the 4 conditions
- Problem-set called Specific Energy, was made available to students from April 17 to April 25 through the course management system
- After 'Specific Energy' homework closed on ANGEL, in-treatment problem-solving test scores were retrieved from the system
- The problem-solving test was administered one week after Specific Energy homework assignment was due, as a part of the course quiz
- After completion of homework and the quiz, all students were given an opportunity to review the most complete course materials, i.e. the treatment that had both segmented multimedia and self-explanation prompts

- Participants completed a survey at the end of the semester
- As an additional source of data, segmented multimedia usage was collected. Reference to the links was tracked through the course management system to examine the extent, to which the treatment materials were actually utilized.
- In-treatment problem-solving performance, problem-solving test scores, survey and multimedia usage were analyzed

Pilot Findings

Table **3.1** shows the scores for the in-treatment problem-solving test.

Table **3.1.** Means and Standard Deviations for Performance on In-treatment Problem-Solving Test

Group	n	In-Treatment Problem Solving	
		M[a]	SD
Segmented Multimedia Worked Examples and Self-Explanation Prompts	17	12.00	3.18
Segmented Multimedia Worked Examples	17	13.76	4.04
Self-Explanation Prompts	16	12.50	4.13
Non-Segmented Multimedia Worked Examples	17	13.82	2.48
Total	67	13.03	3.53

Note a. Maximum score = 21; $F(3, 63) = 1.15$, p =.34

Table **3.2** shows the scores for the delayed problem-solving test. The highest score achieved by the students was 3.

Table 3.2. Means and Standard Deviations for Performance on Delayed Problem-Solving Test

Group	n	Delayed Problem Solving	
		M[a]	SD
Segmented Multimedia Worked Examples and Self-Explanation Prompts	17	2.12	.60
Segmented Multimedia Worked Examples	17	1.94	.74
Self-Explanation Prompts	16	2.25	.68
Non-Segmented Multimedia Worked Examples	17	1.94	.75
Total	67	2.06	.69

Note a. Maximum score = 4; $F(3, 63) = .76$, $p = .52$

There were also several important findings from the survey that influenced the design of the main study. First of all, many students felt that the number of self-explanation boxes that they were required to type in was excessive. One of the explanations for that finding could be that since self-explanation prompts were supporting each individual step in the problem, some of them were essentially the same. At the same time, some of the steps were not very rich in concepts, which resulted in relatively simple self-explanations that could have been perceived as redundant by some of the learners. Secondly, while some of the students perceived self-explanations excessive, others noted that they were very helpful in making them think about the concepts behind the problem-solving procedures. Thirdly, most students perceived multimedia examples as useful, regardless of whether they were exposed to segmented multimedia within the steps or the complete multimedia example in the beginning of the homework.

Multimedia worked example usage data was collected using ANGEL, which recorded hits on the multimedia example links and included the user id, number of hits on each link, and time when each link was accessed. One of the unexpected findings from video usage data was a relatively low number of hits on multimedia segments, as compared to the complete multimedia example. There are several possible explanations for this finding: 1) students might have kept the complete video open and referred to it during the whole problem-solving process; 2) students could have felt that they remembered all relevant information from the complete multimedia example and didn't feel the need to refer to individual segments; or 3) students might have overlooked the individual segments because they were not appropriately labeled.

Implications for the main study

Based on the findings from the pilot, the following changes were made in the main study:

1. The number of self-explanation prompts was decreased to 4 prompts total to address the concern that students would be overwhelmed with the number of self-explanations, which could result in lower quality of self-explanations.
2. The structure of self-explanation prompts was revised to promote more accurate connections among the concepts and to solicit higher quality of self-explanations.
3. Multimedia examples were revised to provide better support for making analogical inferences between the worked example and the practice problems and to promote better conceptual understanding.
4. Placement of segmented multimedia examples was reviewed to provide more explicit connections between the worked example and the practice problems.
5. Labeling for segmented multimedia links was changed to more precise: from "Click here to view the supporting video" to "Click to review the step".

6. Analytical homework problem-solving test items were changed from open-ended to multiple-choice, as it was discovered that the open-ended items could not be reliably scored by the course management system, and some of them had to be re-scored manually.
7. Since the main purpose of the project was to examine the effectiveness of different types of multimedia worked examples, self-explanation prompts-only treatment was removed to better fit the purposes of the main study and the overall project.
8. A new example-problem-set was developed for the main study for the following reasons: (1) this problem-set had to be developed to support one of the difficult concepts of the course; (2) this problem-set was scheduled to be delivered first, based on the course schedule, and it was decided to use this set for the controlled study to ensure clean data.

Main Study

Participants

The participants of the main study were 62 engineering students enrolled in a 300-level Civil Engineering course. For all participants this is a required course for their major. While typically students at this course level are only beginning to specialize in civil engineering, and are assumed not to have substantial knowledge of engineering principles, it was expected that those who enrolled in the course had solid knowledge of algebra and physics. Due to the specifics of engineering student population, and the university demographics, the majority of the participants were white male students around the ages 19-22.

Although treatments and instruments were delivered as a part of the course assignments, only students who had signed the consent form to give permission to use their assignments and grades for the research project were included in the study. Students received extra credit for participation in the end-of-semester survey that was a part of a larger study and therefore was not included in the analysis. Out of 65 participants who signed the consent form, and completed the treatment and the conceptual knowledge test, only 62 participants completed the problem solving test and were included in the data analysis.

Research Design

The independent variable used in the main study was segmented multimedia worked examples with and without self-explanation prompts. The three dependent variables were problem-solving, conceptual knowledge, and quality of self-explanations. See Figure **3.2** for the diagram of the research design.

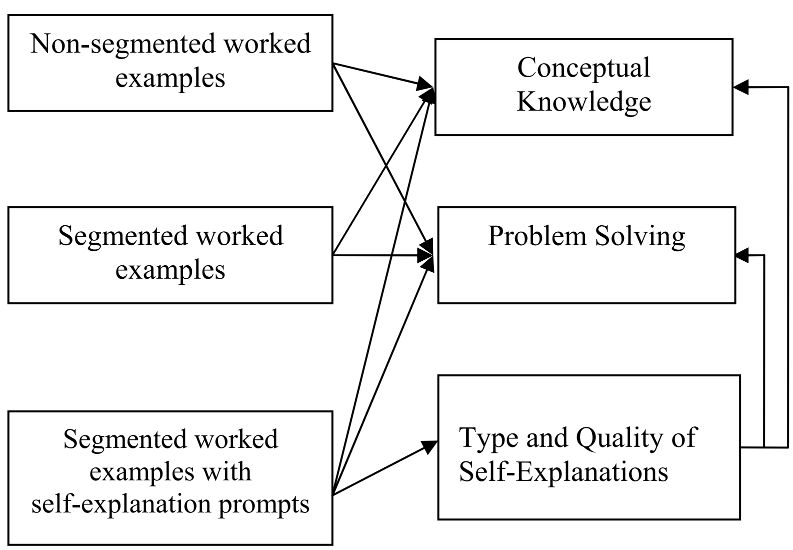


Figure **3.2.** Research Design

Materials

The content of the study came from an undergraduate 300-level course in fluid mechanics. The course was designed to cover fundamental fluid mechanics topics that are used in civil engineering. Some of the main topic areas included: fluid statics, fluids in motion, fluid kinematics, and dimensional analysis. The main topic areas were divided into individual topics, for example, fluids in motion include Newton's 2nd law, Bernoulli equation, energy and hydraulic grade lines, and energy equation. Course activities included two weekly lectures, seven homework assignments, and seven in-class quizzes. Homework assignments and quizzes asked students to solve fluid mechanics problems relevant to the topics presented in class. In Spring 2009 two multimedia example-problem sets were developed to support problem-solving in homework assignments. For Fall 2009 a third problem-set was developed. This problem-set was delivered first in the semester, and it is the one that was chosen to test the effects of the treatments. This example-problem-set was designed to support understanding of energy and hydraulic grade lines in a fluid mechanics course. The content was selected because it was determined by the course instructor that, over the previous semesters, students were consistently having difficulty with both physical and analytical principles involved in this topic. It was also determined that the complexity of the concepts was sufficient to trigger meaningful self-explaining activities.

Practice problems used in the example-problem-set were adapted from the textbook by Crowe et al. (2005) “Engineering Fluid Mechanics 8/E”. It was a different textbook from the one used in class, to make sure students could not check their answers against the answer key typically provided by all textbooks. This was especially important because students were working on the problem-set as a part of their homework.

Each problem-set consisted of a multimedia-based worked example on a fluids mechanics topic and 3 practice problems. The worked example consisted of a number of instructional explanations of the topic, which included drawing learners' attention to various aspects of the physical phenomenon and conceptualizing of analytical information, i.e. explaining the equations, as well as a complete solution procedure for an example problem divided into 4 steps and two conceptual and procedural explanations for more difficult problems. The instructional explanations were delivered as audio over the multimedia presentation, drawing over the pictures and diagrams, and writing in the white space within presentation frames.

All practice problems belonged to the same problem type and used the same concept, Bernoulli principle, but they were presented in the order of increased difficulty: the first practice problem had a very similar structure as the worked example, but different surface features; the second problem employed the same concepts but had a different structure; and the third problem had a different structure and one additional concept, using a manometer to measure pressure. The practice problems consisted of an average of 4 steps. Each step had a procedural label, e.g. 'evaluate the Bernoulli equation', or 'assess pressure changes'. As participants went through the problem-set, they had to type in their answers to the analytical questions, which asked to either calculate a certain parameter of the problem or compare two parameters. Some of the steps required more than one analytical question, and the total number of analytical questions was 26: Problem 1 – 8 items; Problem 2 – 9 items; and Problem 3 – 9 items. The answers were recorded into the course management system (ANGEL) and used for further analysis.

Each problem was built as a separate quiz in the course management system. Once the participants finished one problem, they clicked 'Submit' button, and they were taken to the next problem. Participants could pace themselves within the problem, but after submitting all the answers to one problem, they could not go back. Students could use the equation help sheet, as they went through the problem solving process.

Treatments

To examine the effectiveness of segmented multimedia worked examples and self-explanation prompts, participants were randomly assigned to one of the three treatment options:

- Non-segmented multimedia worked example (NS-MWE)
- Segmented multimedia worked example (S-MWE)
- Segmented multimedia worked example enhanced with self-explanation prompts (S-MWE-SE)

Non-segmented multimedia worked example (NS-MWE)

In this condition, the complete multimedia example was presented after the problem statements and each step. The complete example lasted 17 min 55 s. To find the information related to each step, participants had to scroll through the multimedia presentation. This set-up was necessary to ensure the separation of the worked example from the practice problems to recreate a typical classroom situation. Such set-up was also considered the least effective in the research literature (Renkl, 2005).

Figure **3.3** demonstrates the developed NS-MWE in the browser window:

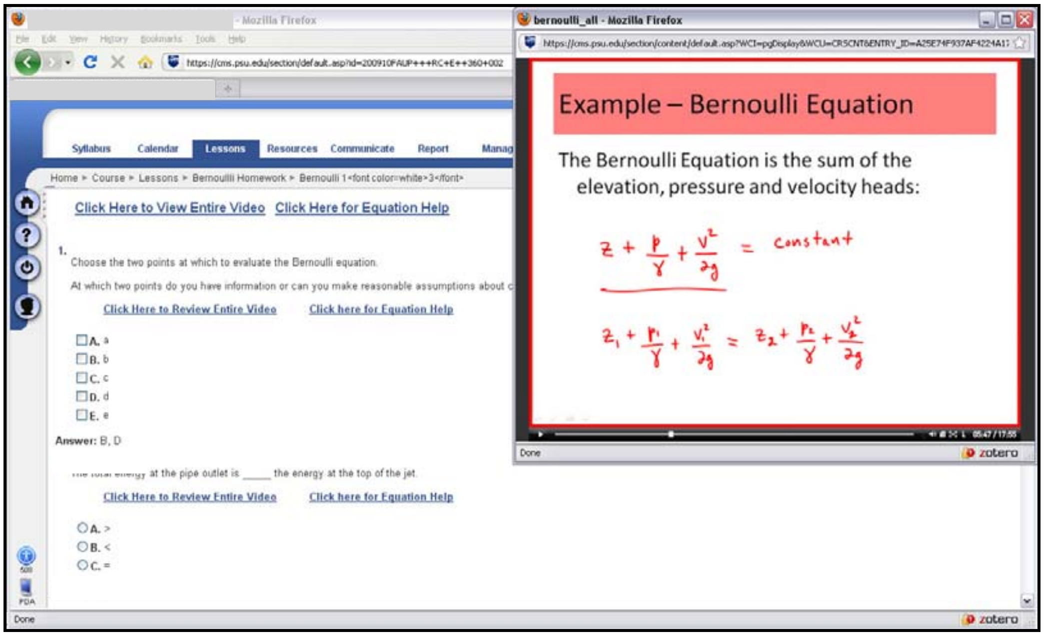


Figure 3.3. Non-Segmented Multimedia Worked Example Screenshot

Segmented multimedia worked example (S-MWE)

For this condition, in addition to the previous set-up, the multimedia worked example was segmented into 9 chunks. The first 3 segments presented conceptual information regarding the problem:

- Segment 1 presented the information about the physical phenomenon and provided real-world examples;
- Segment 2 provided a conceptual explanation of the equations, used for solving the problem;
- Segments 3 and 4 presented conceptual and procedural information relevant for Problems 2 and 3;
- Segment 5 provided the general set-up of the example problem;

- Segments 6-9 present the demonstration of the four steps that are necessary to take to solve Problem 1 and partially Problem 3.

The worked example was linked to each practice problem through individual steps. Participants could also view the complete worked example before each problem. In the beginning of the example-problem-set, participants were instructed to review the segments as they solved for individual steps to make sure that learners were aware of the availability of the instructional supports. Table **3.3** provides duration times and summarizes the content of the complete example and multimedia segments.

Table 3.3. Duration and Content of Multimedia Segments

	Duration	Content
Multimedia Segment		
Complete	17 min 55 s	Full (also includes the real-world example and problem outline)
Segment 1	2 min 23 s	Real-world example
Segment 2	3 min 37 s	Conceptual: explanation of terms of the equation and how they refer to physical principles
Segment 3	3 min 55 s	Conceptual and Procedural Example - manometer
Segment 4	1 min 52 s	Conceptual Example - stagnation tube vs. manometer
Segment 5	1 min 16 s	Problem outline
Segment 6	1 min 10 s	Step 1 - application of Bernoulli equation to 2 points of interest
Segment 7	1 min 17 s	Step 2 - assessment of pressure components (heads) at both points
Segment 8	1 min	Step 3 - solving Bernoulli equation for unknowns
Segment 9	1min 17 s	Step 4 - evaluating the flow rate equation

Table **3.4** demonstrates how segments were linked to each problem: solution steps for Problem 1 were linked in the same order they were presented in the multimedia worked example along with the explanation of terms of the equation and how they referred to physical principles; at the same time, Problems 2 and 3 were mostly supported with additional conceptual and procedural examples explaining how to handle manometer and stagnation tube vs. manometer problems.

Table 3.4. Segment Linkage by In-Treatment Problem Items

	Problem 1					Problem 2					Problem 3							Total links
Item	1	2	3-5	6	7-9	10	11	12-14	15	16-17	18	19-20	21-22	23	24	25	26	
Segment Number																		
2		x		x					x		x				x			5
3							x						x					3
4						x		x		x		x		x		x		10
6	x																	1
7			x															3
8					x													3
9																	x	1

Figure **3.4** demonstrates the developed S-MWE in the browser window:

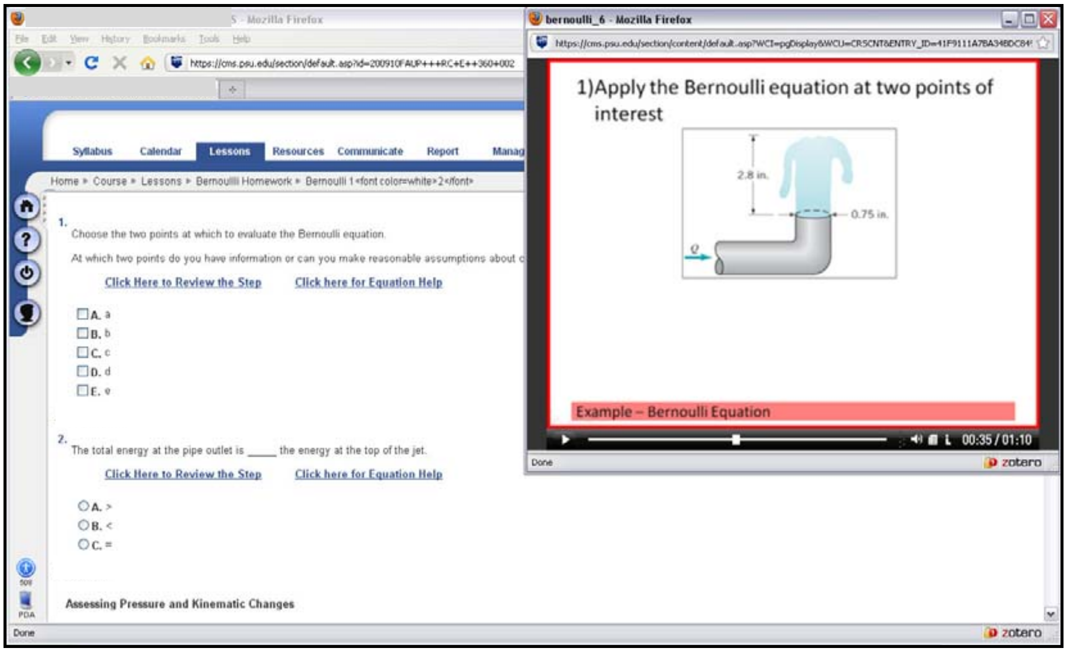


Figure 3.4. Segmented Multimedia Worked Example Screenshot

Segmented multimedia worked example enhanced with self-explanation prompts (S-MWE-SE)

For this condition, in addition to the set-up for S-MWE, participants were prompted to explain their reasoning behind their solution. Self-explanation prompts presented an open-ended question which was located on the same screen as the calculation questions at the places that were indicated by the course instructor as conceptually difficult. There were a total of 5 self-explanations in the problem-set: Problem 1 – 2 self-explanations; Problem 2 – 1 self-explanation; and Problem 3 – 2 self-explanations. The purpose of the prompts was to help participants connect the analytical information with the conceptual information through the self-monitoring process and apply it for the solution of the step. Participants are asked to type in their self-explanations in a text box. The placement and content of self-explanations was determined by the subject-matter expert and the researcher in order to support the most difficult concepts within these problems.

Each self-explanation prompt was placed after a calculation / evaluation question to support that question. For a complete list of self-explanation prompts and preceding calculation questions see Table **3.5**.

Table **3.5**. Calculation Questions and Corresponding Self-Explanation Prompts (SEP) by Problem

Problem Number	SEP Number	Calculation Question	Self-Explanation Prompt
1	1	Choose the two points at which to evaluate the Bernoulli equation.	Please explain why you chose these two locations.
	2	How does the pressure at the top of the jet compare to the pressure at the outlet of the pipe? [<, >, =]	Justify your selection for pressure comparison in the previous question.
2	3	1) In case 1, hA [<, >, =] hB. 2) In case 2, hC [<, >, =] hD.	Justify your selections for both cases 1 and 2 in the previous two questions.
3	4	1) Identify all of the components of total energy that are being ‘measured’ at station 1. 2) Identify all of the components of total energy that are being 'measured' at station 2	Justify your selection of energy components for both stations 1 and 2 in the previous two questions.
	5	Given that V2 = 2V1, the pressure head at station 1 is [<, >, =] the pressure head at station 2.	Justify your selection for the pressure head assessment in the previous question.

Figure **3.5** demonstrates the developed S-MWE-SE in the browser window:

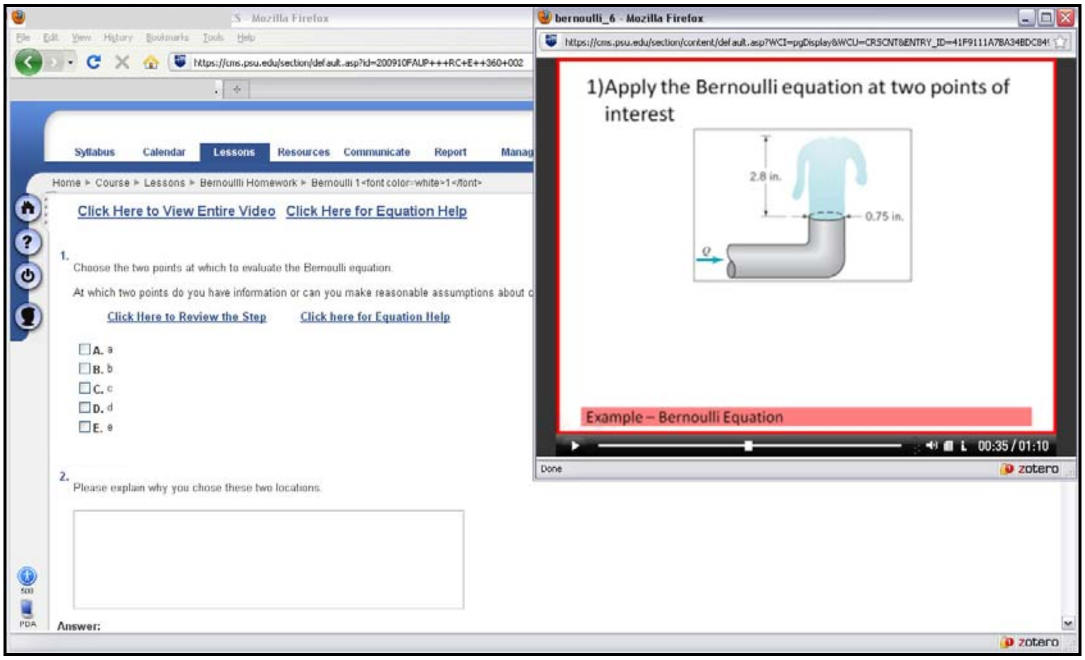


Figure 3.5. Segmented Multimedia Worked Example Enhanced with Self-explanation Prompts Screenshot

Table **3.6** summarizes the differences between treatments in terms of various elements that were either included or excluded from each treatment.

Table 3.6. Summary of Differences in Treatment Design

	Treatment Elements			
	Complete Multimedia Worked Example	7 Linked Multimedia Worked Example Segments	5 Self-Explanation Prompts	3 Practice Problems Divided into Steps
Treatment				
Non-segmented multimedia worked example (NS-MWE)	Yes	No	No	Yes
Segmented multimedia worked example (S-MWE)	Yes	Yes	No	Yes
Segmented multimedia				

worked example enhanced with self-explanation prompts (S-MWE-SE)	Yes	Yes	Yes	Yes

Table **3.7** summarizes the differences in treatment paths intended by the design. However, individual learners could choose to solve the steps within one problem in any order, as well as access or not access the multimedia segments available to them.

Table 3.7. Differences in Designed Paths between Treatments

NS-MWE	**S-MWE**	**S-MWE-SE**
Readme Go to Problem 1 Problem 1 • Read problem statement • Review full example • Solve for step 1 / review full example as needed • Repeat for steps 2-4 • Submit	Readme Go to Problem 1 Problem 1 • Read problem statement • Review full example • Solve for step 1 / view solution for Step 1 • Repeat for steps 2-4 • Submit	Readme Go to Problem 1 Problem 1 • Read problem statement • Review full example • Solve for step 1 / view solution for Step 1 • Type in Self-explanation 1 for Step 1 • Solve for step 2 / view solution for Step 1 • Type in Self-explanation 2 for Step 2 • Solve for steps 3-4 / view solution for Steps 3-4 • Submit
Go to Problem 2 Problem 2 • Read problem statement • Review full example	Go to Problem 2 Problem 2 • Read problem statement • Review full example	Go to Problem 2 Problem 2 • Read problem statement • Review full example • Solve for step 1 / view

• Solve for step 1 / review full example as needed • Repeat for steps 2-4 • Submit	• Solve for step 1 / view supporting solution segment • Repeat for steps 2-4 • Submit	supporting solution segment • Type in Self-explanation 3 for Step 1 • Solve for steps 2-4 / view supporting solution segment • Submit
Go to Problem 3 Problem 3 • Read problem statement • Review full example • Solve for step 1 / review full example as needed • Repeat for steps 2-4 • Submit	Go to Problem 3 Problem 3 • Read problem statement • Review full example • Solve for step 1 / view supporting solution segment • Repeat for steps 2-3 • Solve for step 4 / view solution for Step 4 • Submit	Go to Problem 3 Problem 3 • Read problem statement • Review full example • Solve for step 1 / view supporting solution segment • Type in Self-explanation 4 for Step 1 • Solve for step 2 / view supporting solution segment • Type in Self-explanation 5 for Step 2 • Solve for step 3 / view supporting solution segment • Solve for step 4 / view solution for Step 4 Submit

All participants had to start from the Readme file, which provided instructions how to go through the treatments. From the Readme file, learners were directed to the first problem which started with a problem statement and then suggested to review the full example and equation help. Learners had control over whether to watch the example before or after completing the steps of the problem. Then they were asked to solve the problem using calculation questions for solution steps and different learning supports depending on the treatment (segments of the multimedia example and self-explanation prompts). When the learner considered the problem solved, he/she had to click ‘submit’ to continue to the next problem.

Measurement Instruments

Problem solving of students was measured using a problem-solving test. Conceptual knowledge was measured by a conceptual knowledge test. Type and quality of generated self-explanations was measured by the rubric for the S-MWE-SE group only. These tests were necessary to examine whether different treatments had different effects on the two important engineering competencies: understanding of physical concepts and principles and an ability to apply them in problem-solving.

Additional data, including multimedia worked example usage, time on task, and in-treatment performance, was collected. In-treatment performance, i.e. responses to analytical questions, was collected using treatment logs to further explore the effects of the treatments on the problem-solving process. Time on task and multimedia worked example usage data was collected using ANGEL tracking capabilities to examine how the treatment materials were actually utilized.

Conceptual Knowledge Instrument

Conceptual knowledge instrument was a 10-item multiple choice test, with the minimum score of 0, and maximum score of 10, which probed understanding of the interrelationships among the concepts involved in the problem-set. Five of the items were taken from *Fluid Mechanics Concept Inventory* developed by Martin et al. (2003), used with the permission of the author. The original inventory consisted of 26 items and included a number of concepts typically covered in a fluid mechanics course. The items in the inventory were “constructed so that correct answers to short problems <would> be obtained only by using the correct conceptual framework” (Streveler et al., 2004). Only the items relevant to the concepts used in the study were employed in the conceptual knowledge instrument. The ideas for the rest of the items in the test were taken from two fluid mechanics textbooks: *Fluid Mechanics: Fundamental Applications* by Cengel & Cimbala (2006) and *Engineering Fluid Mechanics 8/E* by Crowe et al. (2005). Neither of these textbooks was used in class, and students were not familiar with the problems from these textbooks.

Content validity of the instrument was established by two engineering faculty, one in civil engineering and one in mechanical engineering by reviewing the questions independently to determine whether they measured the content that was being assessed. As a result of this assessment, one of the problems was removed and replaced with a different problem that was more relevant to the topic. The reliability of the scores was obtained during the main study and was $\alpha = .75$.

A sample item from the test is demonstrated in Figure **3.6** (see Appendix B for the complete instrument):

8. Pitot tubes are placed in two ducts in which air flows as shown below. The density and temperature of the flows are equal. The dynamic (velocity) pressure and the static pressure taps are connected to two manometers. The pressure difference for duct A is 2" of water and that for Duct B is 4" of water. Circle the correct answer for the velocity V_A in duct A relative to the velocity V_B in duct B.

A. V_B equals 2 V_A
B. V_B equals $\sqrt{2}$ V_A
C. V_B equals V_A
D. V_B equals $V_A / \sqrt{2}$
E. V_B equals $V_A /2$

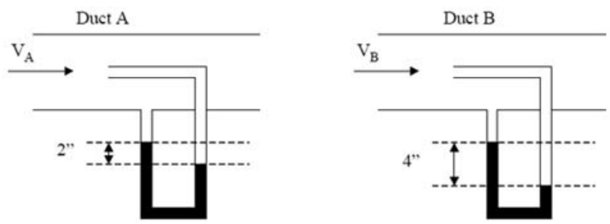
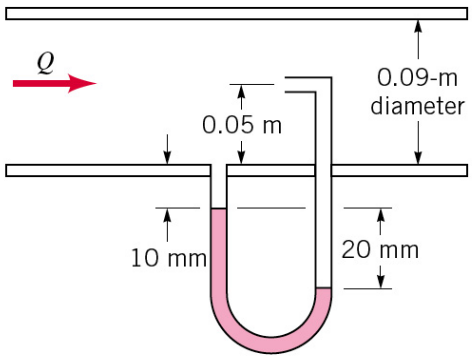


Figure **3.6.** Conceptual Knowledge Test Item

Problem-solving Instrument

The problem-solving test was a partially free-scored test that consisted of 3 problems that were similar to the homework problems. The problems in the test were adapted by the subject matter experts from the textbook by Crowe et al. (2005) *Engineering Fluid Mechanics 8/E*. The subject matter expert determined the relevance and consistency of the items with the treatments to establish the validity of the content.

One of the problems in the test is demonstrated in Figure **3.7** (see Appendix C for the complete instrument):



Oil of specific gravity 0.80 flows in the pressurized constant-diameter pipe shown below. Neglecting head losses, calculate the flow through the pipe. The manometer fluid is water.

Note: Unlike the similar homework problem you did online, here the fluid is oil (not air), so you must take into account the static fluid pressure as it varies with depth.

Figure **3.7.** Problem-Solving Test Item

Problems were scored using the rubric at Table **3.8**. The maximum score was 13. The 3 problems were first graded by the engineering graduate assistant using a rubric that had maximum scores of 1.5, 2, and 8 respectively, with students receiving either complete or partial score, depending on correctness of application of analytical procedures and number of calculation mistakes. However, due to some concerns regarding construct validity and consistency of the first rubric, such as assigning credit for “trying”, “implying”, and “being a good student”, the new rubric, presented in Table 3.8, was developed by the engineering graduate assistant and the researcher, and confirmed by the subject matter expert, that assessed the constructs more consistently. This rubric was scored by the engineering graduate assistant and the researcher with personal identifiers removed from the scores to ensure objectivity. The reliability of the scores obtained from the main study was $\alpha = .70$.

Table 3.8. Problem-Solving Test Scoring Rubric

N	Category	Points
Problem 1		
1	Correctly identify 2 locations	1
2	Correctly explain 2 locations	1
Problem 2		
1	Multiple-choice	1
Problem 3		
1	Pick two appropriate locations	
	Location # 1	1
	Location # 2	1
2	Correctly recognize which Bernoulli equation terms are zero or non-zero	
	Location # 1	1
	Location # 2	1
3	Figure out difference in pressure between 2 points using the manometer equation	
	Correct manometer expression relating p_1 and p_2	1
	Recognizing the need to account for the weight of the fluid in the pipe (in addition to manometer fluid)	1
4	Correct calculation of oil's specific weight (or recognition of the need to do it) $\gamma_{oil} = 0.7_w$ or equivalent (or 0.88_w)	1
5	Carry out calculation to find velocity (V) in the pipe	1
6	Recognize need to use continuity equation Q = VA	1
7	Calculate cross-sectional area of pipe $A = \pi r^2$	1
Total		13

Type and Quality of Self-explanations Measure

There were a total of 5 self-explanation prompts embedded in S-MWE-SE treatment: 2 in problem 1, 1 in problem 2, and 2 in problem 3. There was no existing self-explanation assessment rubric found in the literature that could be applied to this study; therefore a custom rubric was developed to assess the quality of self-explanations. The initial categories for the rubric were adopted from Berthold & Renkl (2009) study, which assessed written self-explanations as one of the variables. According to Berthold & Renkl (2009), “the quality of written self-explanations is a good indicator of the quality of the learning processes” (p. 76). The categories initially emerged from the extensive think-aloud protocol analysis done by Renkl (1997). Since written self-explanations lack some of the qualities of the verbal processes, the categories were adapted to suit the different media:

Principle-based self-explanations: this category refers to the self-explanations that identify the underlying domain principles and elaborate on them.

Rationale-based self-explanations: this category refers to the self-explanations that give reasons why the principle is the way it is.

Incorrect self-explanations in terms of confusion of principles: refers to the self-explanations that demonstrate the learner’s systematic misunderstanding in terms of confusion of principles (Berthold & Renkl, 2009, pp. 76-77).

To develop and finalize the rubric to assess the quality of these self-explanations, self-explanations for prompts 1 and 2 in problem 1 were analyzed. As a result of the analysis, the following criteria for this set of self-explanations emerged: type of self-explanations, quality of self-explanations, and correctness of self-explanations.

Type of self-explanations in this study was the type of knowledge demonstrated: 1) principle-based (conceptual) or 2) procedure-based (problem-solving).

Quality of self-explanations in this study is the extent to which individual learners 1) engaged in identifying and elaborating on the principles or procedures and 2) provided the rationale of why the principle or procedure was used. This definition was partially adopted from (Berthold & Renkl, 2009).

Correctness of self-explanations in this study was the correctness of the content of self-explanations with regards to principles and procedures of the domain of physics and civil engineering.

Table **3.9** specifies the details of the self-explanation assessment rubric.

Table 3.9. Self-Explanation Assessment Rubric

Type	Definition	Incomplete (Level 1)	Complete (Level 2)	Elaborated (Level 3)	Correctness
Principle-based self-explanation	The explanation makes a connection between physical phenomena involved in the process, such as (stream of) water, atmospheric pressure, streamline, jet, manometer, pitot tube	The explanation only partially explains the phenomenon or doesn't make a connection between phenomena	The explanation sufficiently explains the phenomenon and makes a connection between phenomena	The explanation sufficiently explains the phenomenon and makes a connection between phenomena, and provides additional information about the phenomenon, or gives an explicit answer	Yes / No (whether the explanation is correct from the content standpoint)
Procedure-based self-explanation	The explanation refers to completing steps in the problem-solving process, such as cancelling out parts of the equation, evaluating the problem, anticipating next steps etc.	The explanation only partially explains the procedure	The explanation completely explains the procedure	The explanation completely explains the procedure and anticipates the next steps in the solution of the problem.	Yes / No (whether the explanation is correct from the content standpoint)
Combined self-explanation	The explanation has elements of both principle-based and procedure-based explanations.	The explanation involves both physical principles and procedures, but only partially explains one or both of them.	The explanation involves a complete explanation of both principles and procedures.	The explanation involves a complete explanation of both principles and procedures, and either provides additional information	Yes / No (whether the explanation is correct from the content standpoint)

				about the phenomena or anticipates next steps or both.	
Statement (not a self-explanation)	States a fact or repeats the numeric answer, doesn't use causal relationship connectors to justify the answer (because, since, if…then, etc.)	N/A	N/A	N/A	Yes / No (whether the statement is correct from the content standpoint)

An example of a Level 3 Principle-based Self-explanation (from the data):

> *Because the water is no longer experiencing gage pressure, the only pressure being seen at each point then is atmospheric, which, because the 2 points are extremely close relative to the entire height of the atmosphere, is the same at each point.*

An example of a Level 3 Procedure-based Self-explanation (from the data):

> *Because at both points the pressures are equal to zero. And If set the datum at point B, then Zb will be zero. Furthermore the velocity at point D is also Zero. Since the flow rate and area are given, therefore, we can solve for velocity at point B. Last, we can use the bernoulli equation to solve for h.*

Then the coded self-explanations were quantified to obtain feature dimensions that could later be used for examining the relationships between type, quality, and correctness criteria from the assessment rubric in Table 3.9 and acquisition of conceptual knowledge and problem solving performance. However, there were a total of 5 self-explanation prompts, which all possessed the same three criteria. Amershi & Conati (2009) in their study used 3 features for 7 actions, with a total of 21 dimensions. In that case, each seven actions were later individually measured against learning gains, which made the presence of 21 dimensions necessary. This approach did not seem practical for this study, as it was not examining the effects of each individual self-explanation prompt on the learning outcomes, but was looking into the effects of each criterion from the rubric on the outcomes. Thus, in this study, given the research question, all five self-explanations were combined by each of the 3 criteria: type, quality, and correctness, and devise a single score for each dimension.

Quality and correctness criteria had one dimension each and were quantified in a straight-forward manner: quality criterion was a scale from 0 (0 – statement) to 3 (1 – incomplete, 2 – complete, 3 – elaborated), and correctness criterion was a binary variable (0 – incorrect, 1 – correct). However, the type criterion had more than one dimension, so it was quantified differently.

The 4 choices for the type criterion were not completely mutually exclusive. Each type possessed 2 attributes: presence / absence of a principle-based explanation and presence / absence of a procedure based explanation. Thus, a self-explanation classified as principle-based, could be characterized as possessing principle-based explaining and lacking procedure-based explaining. At the same time, a combined self-explanation possessed both qualities, and a statement possessed none. Absence / presence of a quality was expressed by a binary variable 0 / 1. Each attribute of the type criterion as a dimension, therefore possessed the labeling as "principle-based explaining" and "procedure-based explaining". As a result, each self-explanation was characterized by two dimensions related to its type as shown in Figure **3.8**: (0, 0), (1, 0), (0, 1), or (1, 1).

Procedure-based explaining

Principle-based explaining		
	Principle-Based Self-Explanation (0, 1)	Combined Self-Explanation (1, 1)
	Statement (0, 0)	Procedure-Based Self-Explanation (1, 0)

Figure 3.8. Dimensions for Type of Self-Explanations Criterion

These 4 dimensions enable the examination of the relationships between type, quality, and correctness criteria and the learning outcomes: principle-based explaining, procedure-based explaining, quality, and correctness. The scores for each of the five self-explanation prompts (SEP) were combined. Each participant had 5 scores, each with four dimensions. Table **3.10** demonstrates a sample score for a participant. For example, if a participant had 3 principle-based explanations and 2 procedure-based explanations, the quality for two of them was level 1, for two was level 2, and for 1 was level 3, and all of them were correct, the participant had the total score (2, 3, 9, 5).

Table **3.10.** Sample Score for Each of the Dimensions by Participant

	Procedure-based Explaining	Principle-based Explaining	Quality of Explaining	Correctness of Explaining
By Participant	2	3	9	5

Self-explanations were coded by 2 different raters based on the rubric in Table 3.9. Then the codes were quantified, and inter-rater reliability analysis using the Kappa statistic for procedure-based, principle-based, and quality dimensions was performed to determine consistency among raters. Procedure-based explaining had inter-rater reliability Kappa =.873 with $p < 0.001$. Principle-based explaining had inter-rater reliability Kappa =.789 with $p < 0.001$. Quality had inter-rater reliability Kappa =.639 with $p < 0.001$.

After re-examining the codes by the two raters together, it was possible to reach agreement on all codes for principle-based and procedure-based explaining, and it was possible to achieve inter-rater reliability Kappa =. 907 with $p < 0.001$ for the quality dimension. In the cases where agreement was not achieved, a mean score of the two ratings was assigned.

Additional data

In-treatment Performance

The treatment was a part of the homework that consisted of three problems. In-treatment scores were examined to further explain the effects of the treatments on the problem-solving process. To receive a score for each problem, learners had to answer a number of multiple choice analytical questions (either calculation, comparison or evaluative). One item from Problem 2 was removed from scoring because the wording of the question opened the possibility of 2 correct answers. One item from Problem 1 and five items from Problem 3 were rescored to correct scoring issues for checkbox questions that the course management system had. The resulting maximum score was 43. The reliability of the scores obtained in the main study was α = .71.

An example of a calculation question:

Compute V at the end of the pipe _____ m/s

A. *6.4 m/s* **B.** *12.7 m/s* **C.** *20.0 m/s* **D.** *50.9 m/s*

An example of a comparison question:

How does the pressure at the top of the jet compare to the pressure at the outlet of the pipe?

A. = **B.** < **C.** >

An example of an evaluative question (checkbox):

Identify all of the components of total energy that are being 'measured' at station 1.

A. *elevation head* **B.** *velocity head* **C.** *pressure head* **D.** *head loss*

Multimedia Worked Example Access

As an additional source of data, ANGEL information on segmented multimedia usage was collected. References to the links as well as participants time stamps were tracked through the course management system to examine how the treatment materials were actually utilized.

To evaluate individual interactions with the treatment materials, the technique called educational data mining (EDM) was used. Education data mining explores “unique types of data that come from educational settings” (Baker & Yacef, 2009) to discover meaningful patterns that would inform researchers on how students learn in particular settings, and subsequently to enhance teaching and learning (Merceron & Yacef, 2005). These types of data typically come from log files containing user interactions with the learning environment, such as time on task, access to multimedia files, and input and submission of work results into the database (e.g. a course management system). In the current study, this strategy was used to identify common interaction behaviors and explain some of the results from the objective measures that were used in the study: a conceptual knowledge instrument and a problem-solving instrument.

While the materials were structured in the order that the researcher and the subject matter experts considered the most effective, some of the settings allowed user control over some of the interactions, as described earlier in this chapter. In addition, the study intention was to have learners complete all 3 problems in one sitting. For the purpose of obtaining cleaner data, students were only allowed to access each problem once; if they were logged out for some reason, they had to ask the researcher to create a new record. However, the amount of time spent on each problem was only limited by the course management system that would go into a time-out every 50 minutes of inactivity. There were no direct means used to access the time spent off-task, but some inferences on this aspect could be made using other log files.

Another factor controlled by the users was the time between completion of the treatment and the problem solving test: while the conceptual knowledge test was completed online immediately after the treatment, the problem solving test was completed in class one week after the treatments were open online, i.e. the treatments could be completed from one week to 30 minutes before the problem solving test.

Information regarding the day of completing the treatment was collected to determine whether additional factors, such as possible discussion of treatments by the students before the problem-solving test, could contribute to the results. However, it is important to mention that there were other possible contributing factors that could not be determined from the logs, such as attendance of the teaching assistant's office hours or individual study of the materials before the test.

Time-stamps were collected to examine differences in time on task between different treatments. Time on task was a sum of times for each problem, which was counted as time when each problem was opened and submitted.

K-means clustering was used to identify patterns of multimedia access behavior. Clustering is one of the techniques that are widely used in EDM to discover behavior patterns (Baker & Yacef, 2009). Cluster analysis is a classification method that is used to group objects possessing similar features into distinctive categories, or clusters, thus organizing the data into a meaningful structure. K-means algorithm is a nonhierarchical clustering method that was introduced by McQueen (1967). This algorithm is iterative in nature, and is based on randomly selected initial partitioning (Anderberg, 1973). It begins with an initial partition of the data set into a pre-defined number of clusters; then it allocates each object to the nearest cluster; then recalculates cluster centers; and then iterates the previous two steps until cluster centers are no longer changing (Anderberg, 1973). The main disadvantage of *k*-means clustering is that the number of clusters has to be specified in advance, and the initial cluster centers are assigned randomly. To overcome this disadvantage, some researchers suggest repeating the procedure with a different number of clusters, i.e. $k = 2, 3, 4$ (Amershi & Conati, 2009), or repeating the procedure with the same number of clusters multiple times to confirm cluster centers and allocations. In this study, a two-stage clustering procedure was used which consisted of average linkage and k-means algorithms. Average linkage is a hierarchical procedure that is based on the average Euclidean distance between objects in each cluster. Combining average linkage with *k*-mean clustering was found a superior procedure to *k*-means clustering alone (Milligan, 1980). The first part of the two-stage procedure described by Avramov (2005) uses the combination of multiple correlation parameter R^2, *pseudo-F*, and *pseudo-*t^2 statistics to select the number of clusters and initial cluster centers. R^2 in this case is defined as the proportion of variance explained by the current number of clusters; the *pseudo-F* statistic is defined as the ratio of between-cluster and within-cluster variances; and the *pseudo-*t^2 statistic is defined as "the ratio of the sum of squared deviations from the cluster center within each of the two clusters that will be merged and the sum of square errors within the resultant cluster" (Avramov, 2005, p. 34). The procedure starts with having as many clusters as observations, and gradually merges all observations into one cluster. A possible solution is indicated by a sudden drop in R^2 which indicates that two very different clusters were merged, and by a local peak in *pseudo-F* followed by a local peak in *pseudo-*t^2. After the number of clusters and the initial cluster centers were identified using this average linkage procedure, *k*-means clustering was applied to determine final cluster centers and cluster allocations. A one-way ANOVA was performed to determine whether different multimedia access patterns, i.e. different clusters, were associated with different learning outcomes.

Procedures

Prior to intervention

In the beginning of the semester, the instructor explained to the students that there would be a number of innovative learning activities carried out and assessed in the course for the purpose of improving the instruction. Students were randomly assigned to three treatment groups using a built-in random group generator in ANGEL. The random assignment was automatically connected to the students' userID.

During the intervention

The homework containing the example-problem-set was made available on ANGEL after the instructor presented the topic during the lecture. The instructor announced the homework in class, and sent an email, reminding the students to complete it within one week. Students could complete the homework at any time during that week, but they had to do it in one sitting. Several days before the homework was opened, the course instructor sent an email to class, explaining how to students should proceed through the homework and reminding them that they would need the headphones to use the multimedia worked examples. Students were also asked to work individually.

When participants opened the homework, they first read the instructions how to proceed through the homework, and then they were exposed to one of the three treatments. Once the homework was accessed, it had to be completed. The completion of the homework took an average of 1hour 25 min.

The timer was set for the homework to close after the one-week deadline. The course instructor sent a reminder to the class one day before the homework was due.

After the intervention

The conceptual knowledge test was opened for completion as the fourth component of the homework. Students could not access the test before completing all three homework problems. Though the answers to the questions could not be easily found in the textbook, a 20-minute timer was set on the test to prevent extensive searching for information.

The problem-solving test was administered the same day the homework was due as a part of an in-class quiz. The answers were graded by the teaching assistant, and the copies of the quizzes were provided to the researcher. The researcher observed the teaching assistant grading some of the quizzes and also received a grading rubric for the quiz.

Since the study was conducted as a part of a major classroom research initiative, the consent forms, asking for release of the results of students' work, were collected at the end of the semester together with other project paperwork for the purpose of better class time management. The researcher and the principal investigator on a major educational engineering research project of which this study was a part went to one of the class sessions in the end of the semester. They explained the purpose of the study and participants' rights. Then they distributed and collected informed consent forms (Appendix A). No data was analyzed before the consent forms were collected. Personal identifiers were removed from the paper copies of the quiz and replaced by numeric codes. All personal data was also removed from the test scores retrieved from ANGEL and replaced by numeric codes as well.

Data Analysis

After identifying the levels of self-regulated learning skills, descriptive statistics, including means and standard deviations of dependent variables were calculated using statistics software (SPSS).

A one-way analysis of variance (ANOVA) instead of multivariate analysis of variance (MANOVA) was used to determine the differences among the experimental groups for both dependent measures, as the dependent measures were not significantly correlated.

Written self-explanations were examined using both qualitative and quantitative means of analysis: 1) initially self-explanations were categorized using the assessment rubric (Table 3.9); 2) categorized self-explanations were quantified; 3) relationships between various self-explanation dimensions and acquisition of conceptual knowledge, and problem-solving performance were examined using scatter plots, correlation analysis, and multiple regression analysis.

Multimedia segment access data was organized into individual profiles. These profiles were quantified and analyzed using *k*-means cluster analysis to identify groups with similar multimedia access patterns. These clusters were analyzed using one-way ANOVA for the conceptual knowledge and problem solving tests as dependent variables to determine which patterns of interaction behaviors were more effective or less effective, as well as to develop better understanding which interaction behaviors are more characteristic for high and low achieving students.

Chapter 4

RESULTS

Introduction

The purpose of this study was to investigate the effects of different types of multimedia worked examples, operationalized as non-segmented (NS-MWE) and segmented multimedia worked examples (S-MWE), and segmented multimedia worked examples enhanced with self-explanation prompts (S-MWE-SE), on acquisition of conceptual knowledge and problem-solving performance in an undergraduate engineering course.

The study also examined whether self-explanations influenced acquisition of conceptual knowledge and problem-solving performance. Learner generated written self-explanations were examined to explore whether there was an interrelationship between the quality of self-explanations, acquisition of conceptual knowledge, and problem-solving performance.

This chapter reports the results of statistical and qualitative analysis to answer the research questions proposed in Chapter 1.

Acquisition of Conceptual Knowledge and Problem Solving Performance

To answer two main research questions: 1) whether different types of multimedia worked examples influenced acquisition of conceptual knowledge; and 2) whether different types of multimedia worked examples influenced problem solving performance, the differences in the conceptual knowledge scores and problem solving scores between the three treatments were examined.

Initially, it was planned to use the MANOVA test to examine the effect of the treatments on both dependent variables at the same time. However, the correlation analysis revealed that there was no significant relationship between conceptual knowledge scores and problem-solving scores: Pearson's $r = .282$, correlation insignificant at $p < .05$ level (2-tailed). Therefore, use of MANOVA was not justified. Instead, ANOVA was used to examine the first two research questions.

Equality of variances and normality of distribution assumptions for ANOVA were examined. Levene's test for equality of variances did not reveal any significant differences between any of the compared groups on either conceptual knowledge test, problem solving test, time on task, or in-treatment problem-solving performance. As far as normality of distribution assumption, *F*-test is a robust test against departures from normality, provided that the departure from normality is not extreme (Kutner et al., 2005). Since kurtosis of a normal distribution is 3, and skewness is 0, kurtosis and skewness for the dependent variables were examined to determine the departure for normality, and the absolute value of kurtosis did not exceed 4, and the absolute value of skewness did not exceed 1 for either conceptual knowledge test, problem solving test, time on task, or in-treatment problem-solving performance.

The descriptive results of the conceptual knowledge test are reflected in Table **4.1.**

Table **4.1**. Means and Standard Deviations for Performance on Conceptual Knowledge Test

		Conceptual Knowledge	
Group	n	M[a]	SD
Segmented Multimedia Worked Examples and Self-Explanation Prompts (S-MWE-SE)	19	6.32	2.58
Segmented Multimedia Worked Examples (S-MWE)	23	6.26	2.52
Non-Segmented Multimedia Worked Examples (NS-MWE)	20	6.25	2.57
Total	62	6.27	2.56

Note a. Maximum score = 10

The ANOVA results indicated no statistically significant differences between any of the groups: $F(2, 59) = .00$, $p = .99$.

The descriptive results for the problem solving performance test are shown in Table **4.2**.

Table **4.2**. Means and Standard Deviations for Performance on Problem Solving Test

		Problem Solving	
Group	n	M[a]	SD
Segmented Multimedia Worked Examples and Self-Explanation Prompts (S-MWE-SE)	19	7.32	2.41
Segmented Multimedia Worked Examples (S-MWE)	23	7.57	2.44
Non-Segmented Multimedia Worked Examples (NS-MWE)	20	8.40	2.50
Total	62	7.76	2.49

Note a. Maximum score = 13

The ANOVA results indicated no statistically significant differences between any of the groups: $F(2, 59) = 1.04$, $p = .36$.

Self-Explanation Prompts

To answer the third research question: whether there was an interrelationship between the type and quality of self-explanations, acquisition of conceptual knowledge, and problem solving performance, two phases of analysis were conducted: first, students' written self-explanations were evaluated and categorized into several dimensions, based on the existing research and the themes emerging from our present set of data; and second, possible interrelationships were examined among the type and quality of generated self-explanations and performance on the two tests to determine whether certain types of self-explanation behaviors were related to certain learning outcomes.

Phase 1: Categorization of Self-Explanations

Ninety-five self-explanations, 5 per student, generated by 19 participants from the segmented multimedia worked examples with self-explanation prompts group (S-MWE-SE), were analyzed. As described in Table 3.4 of Chapter 3, qualitative characteristics of these self-explanations were analyzed based on the following criteria: type of self-explanations, quality of self-explanations, and correctness of self-explanations. Table **4.3** provides frequency distributions for different types of self-explanations and statements per self-explanation prompt (SEP). The table demonstrates that patterns of self-explanations by type were different depending on a self-explanation prompt (SEP).

Table 4.3. Frequency Distributions for Type of Self-Explanations

	SEP 1		SEP 2		SEP 3		SEP 4		SEP 5		Total	
	n	Percent	n	Percent	n	Percent	n	Percent	n	Percent	n	Percent
Statements	3	3.2%	1	1%	0	0%	9	9.5%	4	4.2%	17	17.9%
Procedure-based	8	8.4%	3	3.2%	5	5.2%	2	2.1%	10	10.5%	28	29.4%
Principle-based	4	4.2%	13	13.7%	7	7.4%	8	8.4%	3	3.2%	35	36.9%
Combined	4	4.2%	2	2.1%	7	7.4%	0	0%	2	2.1%	15	15.8%
Total	19	20%	19	20%	19	20%	19	20%	19	20%	95	100%

Note. SEP – Self-Explanation Prompt

Table **4.4** demonstrates the distribution of self-explanations by type by participant. The results indicate that all participants engaged in self-explaining rather than stating for at least 2 self-explanation prompts and 7 participants (almost 37%) engaged in self-explaining for all five prompts. It also demonstrates that almost all participants (95%) employed both procedure-based and principle-based explaining.

Table 4.4. Distribution of Self-Explanation by Type by Participant

Participant	Statement	Procedure-Based	Principle-Based	Combined
4	0	2	3	0
14	1	2	1	1
15	2	1	2	0
16	1	1	1	2
19	2	1	2	0
22	2	1	2	0

23	0	3	2	0
29	0	0	3	2
35	1	2	2	0
36	0	4	1	0
37	0	3	2	0
40	0	0	4	1
41	0	0	2	3
49	1	1	3	0
58	1	1	1	2
59	1	1	3	0
62	3	2	0	0
64	1	2	0	2
66	1	1	1	2
Total	17	28	35	15

Then self-explanations were quantified, and each self-explanation was assigned a 4-dimensional score, as described in Chapter 3. That score indicated the level of procedure-based explaining, principle-based explaining, quality, and correctness exhibited by each participant. The 4-dimensional scores for all 5 self-explanation prompts were added to obtain a single 4-dimensional score for each participant. Table **4.5** provides means and standard deviations for each of the four dimensions. The results demonstrate that participants used slightly more principle-based explaining than procedure-based explaining, and their average quality of self-explaining was adequate for their level of expertise, but not particularly high. The level of correctness was also adequate.

Table **4.5**. Means and Standard Deviations for Four Self-Explanation Dimensions (n = 19)

Self-Explanation Dimension	Dimension Score	
	M	SD
Procedure-Based Explaining [a]	2.21	.98
Principle-Based Explaining [a]	2.68	1.25
Quality of Self-Explanation [b]	7.42	2.30
Correctness of Self-Explanation [a]	3.58	.96

Note a. Maximum score = 5
Note b. Maximum score = 15

Table **4.6** demonstrates the distribution of two types of self-explaining by quality and correctness. Statements were not coded for the type of explaining, therefore they were excluded from this table. The results demonstrate that the levels of quality and correctness of both types of self-explaining were relatively similar.

Table **4.6.** Distribution of Procedure-Based and Principle-Based Explaining by Quality and Correctness

Quality	Procedure-Based Explaining		Principle-Based Explaining	
	Correct	Incorrect	Correct	Incorrect
1 – 1.5	10	2	17	0
2 – 2.5	19	5	17	8
3	5	0	3	0

Phase 2: Relationships between Self-Explanation Dimensions and Performance on the Outcome Measures

Self-explanation dimensions were examined to determine whether any relationships existed between each or all of the dimensions and performance outcomes. Table **4.7** summarizes correlations between self-explanation dimensions and performance on the tests.

Table **4.7.** Correlations between Self-Explanation Dimensions and Test Performance (n = 19)

	Procedure-based	Principle-based	Quality	Correctness	Conceptual Knowledge	Problem Solving
Procedure-based	–	-.079	.119	.278	.059	.039
Principle-based		–	.735**	.255	.082	.436
Quality			–	.560*	-.082	.461*
Correctness				–	.440	.221
Conceptual Knowledge					–	.282
Problem Solving						–

Note. ** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).

To explore the relationships between the self-explanation dimensions and learning outcomes that did not achieve significance through linear correlation analysis, scatter plots were produced with each of the dimensions as independent variables and conceptual knowledge and problem solving performance measures as dependent variables, and the line was fitted to determine the nature of the relationship. Figures **4.1-4.4** demonstrate the relationships between the conceptual knowledge and problem solving test and those self-explanation variables where it was possible to determine the nature of the relationship. As expected from the correlation analysis results, there was a linear relationship between correctness of self-explanations and conceptual knowledge (Figure **4.1**), as well as principle-based explaining and problem solving (Figure **4.2**). In addition, as evidenced in the scatter plots in Figure **4.3** and Figure **4.4**, there were possible non-linear relationships between procedure-based explaining and conceptual knowledge (Figure **4.3**) and principle-based explaining and conceptual knowledge (Figure **4.4**).

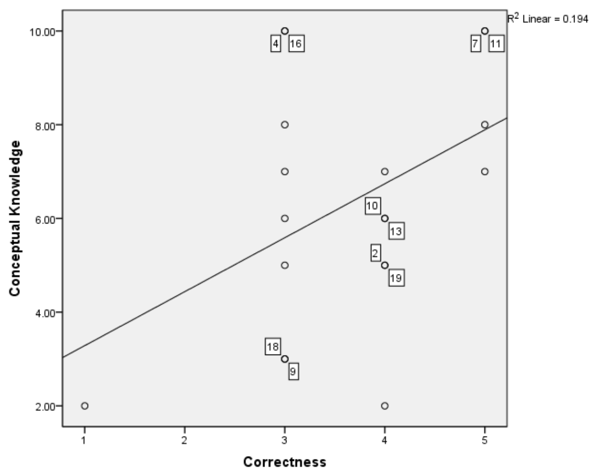


Figure **4.1.** Correctness vs. Conceptual Knowledge

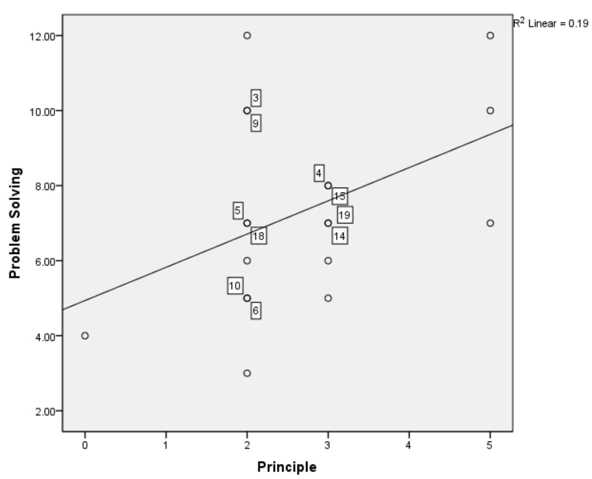


Figure **4.2.** Principle-based Explaining vs. Problem-Solving Performance

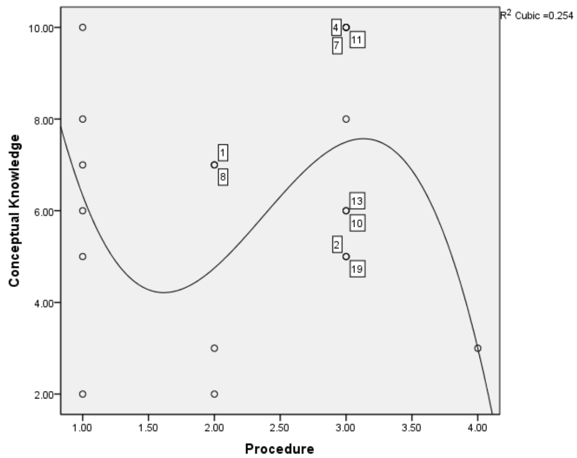


Figure 4.3. Procedure-based Explaining vs. Conceptual Knowledge

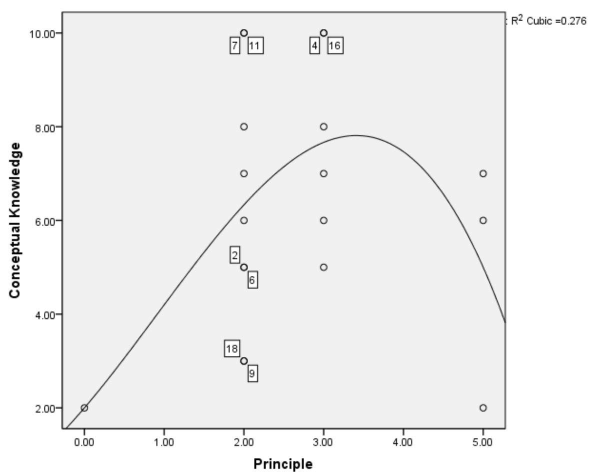


Figure 4.4. Principle-based Explaining vs. Conceptual Knowledge

A multiple regression analysis was conducted to determine whether any of the self-explanation dimensions (procedure-based explaining, principle-based explaining, quality, and correctness) were significant predictors of the performance on the tests. The stepwise regression method was selected because the correlation results suggested possible multicollinearity among several of the predictor variables (Kutner et al., 2005). No model could be produced for the conceptual knowledge measure. A model for the problem solving measure is presented in Table

4.8. Quality of self-explanations significantly predicted problem solving scores: *F Change* (1, 17) = 4.59, *Sig. F Change* = .047.

Table **4.8.** Multiple Regression Results for Interrelationship Between SE Dimensions and Problem-Solving Performance (n = 19)

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.461[ab]	.213	.166	2.319	.213	4.594	1	17	.047

Note a. Predictors: (Constant), Quality
Note b. Excluded: Procedure, Principle, Correctness

Differences in self-explaining between 5 self-explanation prompts were also explored. Figure **4.5** represents average levels for type dimensions (procedure-based and principle-based explaining) by self-explanation prompt (SEP). The graphs for SEP1 and SEP 5 indicate the highest amounts of procedure-based explaining, SEP2 and SEP3 indicate the highest amount of principle-based explaining, and the graph for SEP4 indicates the lowest amount of procedure-based explaining and a relatively low level of the principle-based explaining among all SEP-s.

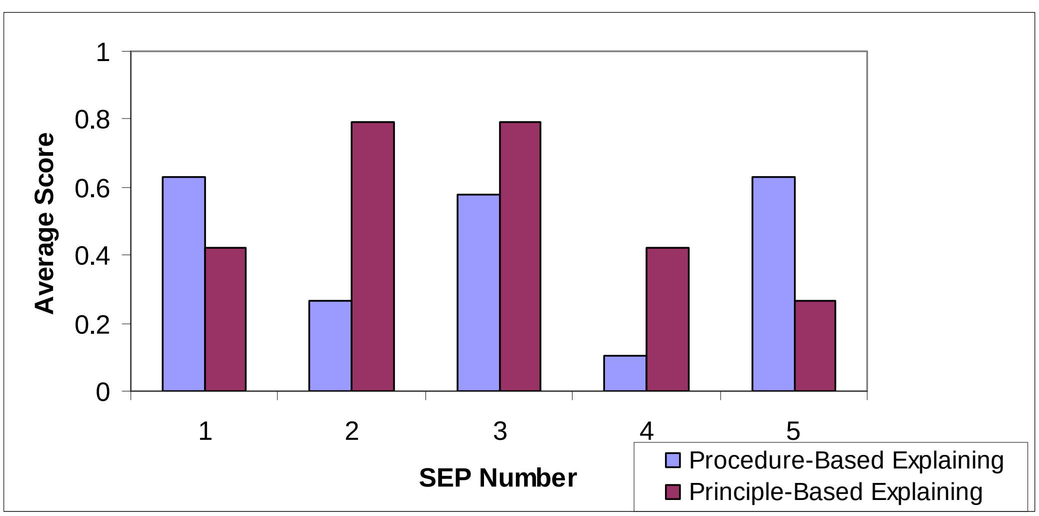


Figure **4.5.** Self-Explanation Levels for Type Dimension by SEP

Figure **4.6** represents average levels for all dimensions by self-explanation prompt (SEP). The graph demonstrates that SEP4 has lowest levels of quality and correctness of self-explaining.

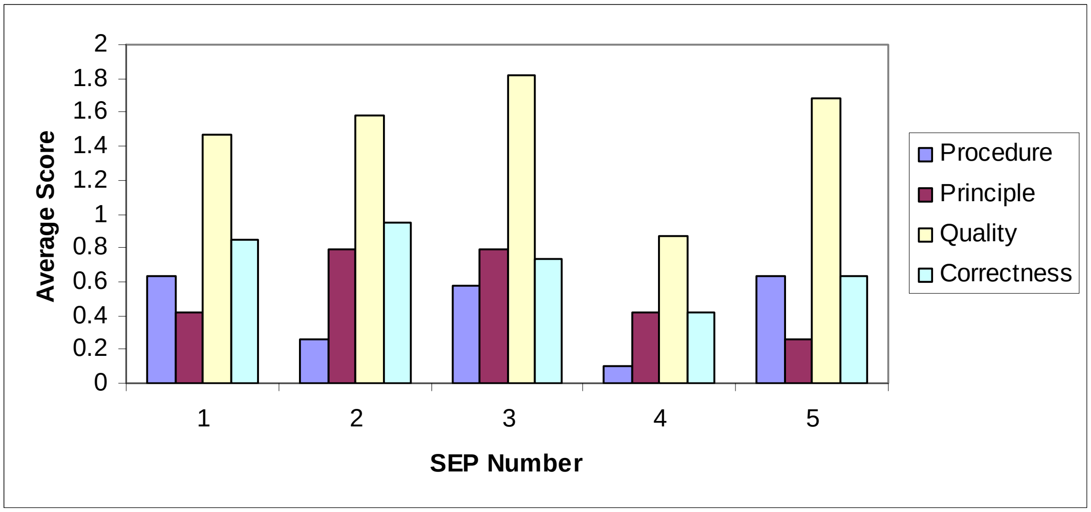


Figure **4.6.** Self-Explanation Levels for All Dimensions by SEP

Usage Data: Learner Interactions with the Environment

Usage Findings

There were three types of additional data that were used to examine learner interactions with the environment:

- treatment scores: user responses for questions within the treatments (43 answers to solution steps and 5 open-ended self-explanations);
- treatment usage logs: time and date when the treatment was started by each user, and time and date when each problem was submitted by each user;
- multimedia usage logs: frequency of access to the multimedia example files by each user

All three types of data were retrieved from the course management system and were used to further explore the research questions.

In-Treatment Problem-Solving Performance

The descriptive results of participants' problem solving performance on the in-treatment problems are reflected in Table **4.9.**

Table **4.9**. Means and Standard Deviations for Problem Solving Performance in Treatment

		In-Treatment Performance	
Group	n	M[a]	SD
Segmented Multimedia Worked Examples and Self-Explanation Prompts (S-MWE-SE)	19	30.37	3.88
Segmented Multimedia Worked Examples (S-MWE)	23	28.61	3.87
Non-Segmented Multimedia Worked Examples (NS-MWE)	20	25.75	5.34
Total	62	28.24	4.46

Note a. Maximum score = 43

The ANOVA results revealed statistically significant differences between the groups: $F(2, 59) = 5.22$, $p = .008$. Tukey *a* post-hoc comparison test revealed significant differences between S-MWE-SE and NS-MWE groups ($p = .006$). No significant differences for in-treatment scores were found at $p = .05$ level between S-MWE-SE and S-MWE or between S-MWE and NS-MWE groups.

Treatment Logs

Treatment logs provided the following data: time participants spent on each problem and date and time when the treatments and the conceptual knowledge test were started and completed. Table **4.10** demonstrates descriptive statistics for the time in minutes that participants spent on the treatment.

Table **4.10**. Means and Standard Deviations for Time on Task

		Time on Task	
Group	n	M[a]	SD
Segmented Multimedia Worked Examples and Self-Explanation Prompts (S-MWE-SE)	19	98.21	58.37
Segmented Multimedia Worked Examples (S-MWE)	23	86.43	36.10
Non-Segmented Multimedia Worked Examples (NS-MWE)	20	71.75	33.44
Total	62	85.47	42.30

Note a. Time in minutes

The ANOVA results indicated no statistically significant differences between the groups: F (2, 59) = 1.86, p = .17. As indicated by the results of the box plots, there was an outlier found for S-MWE-SE group: with the outlier removed the descriptive statistics for this group were M = 88.11, SD = 37.84.

The treatments were open for participants 7 days before the in-class problem solving test. Table **4.11** shows the number of participants completing the treatments from the opening day (October 1, 9 p.m.) to the day of the test (October 8, 3:30 p.m.):

Table 4.11. Number and Frequency Distribution for Treatment Completion

	S-MWE-SE		S-MWE		NS-MWE		Total	
Date [a]	n	Percent	n	Percent	n	Percent	n	Percent
October 2	0	0%	1	4.3%	0	0%	1	1.6%
October 4	1	5.3%	0	0%	1	5%	2	3.2%
October 5	1	5.3%	1	4.3%	0	0%	2	3.2%
October 6	3	15.8%	0	0%	0	0%	3	4.8%
October 2	10	52.6%	11	47.8%	12	60%	33	53.2%
October 7	4	21%	10	43.5%	7	35%	21	33.9%
October 8	19	100%	23	100%	20	100%	62	100%
Total	0	0%	1	4.3%	0	0%	1	1.6%

Note a. No participants completed treatments on October 1 or October 3

These results demonstrate that the majority of participants (87.1%) completed the treatment not earlier than 1 day before the problem-solving test.

Multimedia Example Usage

Table **4.12** presents a summary of descriptive results for the multimedia example usage by treatment group. The results demonstrate that the non-segmented group was accessing the complete multimedia example slightly more frequently, especially compared to the segmented multimedia worked examples and self-explanation prompts group.

Table 4.12. Means and Standard Deviations for Multimedia Access

Group	n	Complete Example		Total Number of Segments		Unique Segments	
		M[a]	SD	M[a]	SD	M[b]	SD
Segmented Multimedia Worked Examples and Self-Explanation Prompts (S-MWE-SE)	19	1.42	1.39	6.89	6.65	3.58	2.71
Segmented Multimedia Worked Examples (S-MWE)	23	1.74	1.21	6.30	4.45	3.84	2.10
Non-Segmented Multimedia Worked Examples (NS-MWE)	20	2.2	.77	N/A	N/A	N/A	N/A
Total	62	1.79	1.12	6.60	5.40	3.70	2.41

Note a. Maximum number of viewings is unlimited
Note b. Maximum number of unique viewings = 7

None of the comparisons revealed significant differences between the groups. The findings from the data will be further elaborated because of the issues discovered in the logs.

Additional Usage Findings

Closer investigation of the treatment log files suggested a high likelihood of a number of participants collaborating on the treatments. This finding was supported by the following evidence: 1) date and time of completing the treatment; 2) treatment scores; and 3) conceptual knowledge score, as the conceptual knowledge test was administered in the same online session, immediately after the completion of the treatment. A total of 17 collaborating participants were discovered: 7 in S-MWE-SE, 6 in S-MWE, and 4 in NS-MWE treatment. There were 2 patterns of collaboration found: 1) two participants assigned to the same treatment working together; and 2) a group of 5-6 participants assigned to different treatments working together.

To evaluate the impact of this collaboration on the results, the following questions were addressed: 1) what is the effect of the treatments on the performance on the tests for non-collaborating participants, that is, if all collaborating participants were removed; and 2) how do the results of collaborating participants compare to the results of non-collaborating participants.

Performance on the Outcome Measures, In-Treatment Scores, and Time on Task for Non-Collaborating Participants

To answer the first question, descriptive statistical tests and ANOVA were conducted for 45 non-collaborating participants for the conceptual knowledge test, problem-solving test, the treatment, and time on the treatment. In addition, a correlational analysis of the self-explanation component was conducted for 12 participants from S-MWE-SE group. Tables **4.13-4.16** provide the descriptive results:

Table **4.13**. Means and Standard Deviations for Performance on Conceptual Knowledge Test

		Conceptual Knowledge	
Group	n	M[a]	SD
Segmented Multimedia Worked Examples and Self-Explanation Prompts (S-MWE-SE)	12	5.58	1.98
Segmented Multimedia Worked Examples (S-MWE)	17	5.47	2.45
Non-Segmented Multimedia Worked Examples (NS-MWE)	16	5.63	2.45
Total	45	5.56	2.29

Note a. Maximum score = 10

The ANOVA results indicated no statistically significant differences between any of the groups: $F(2, 42) = .02, p = .98$.

Table 4.14. Means and Standard Deviations for Performance on Problem Solving Test

		Problem Solving	
Group	n	M[a]	SD
Segmented Multimedia Worked Examples and Self-Explanation Prompts (S-MWE-SE)	12	6.92	2.71
Segmented Multimedia Worked Examples (S-MWE)	17	7.94	2.48
Non-Segmented Multimedia Worked Examples (NS-MWE)	16	8.31	2.44
Total	45	7.72	2.55

Note a. Maximum score = 13

The ANOVA results indicated no statistically significant differences between any of the groups: F (2, 42) = 1.08, p = .35.

Table 4.15. Means and Standard Deviations for Problem Solving Performance in Treatment

		In-Treatment Performance	
Group	n	M[a]	SD
Segmented Multimedia Worked Examples and Self-Explanation Prompts (S-MWE-SE)	12	29.83	4.15
Segmented Multimedia Worked Examples (S-MWE)	17	29.00	4.15
Non-Segmented Multimedia Worked Examples (NS-MWE)	16	24.75	5.72
Total	45	27.86	4.68

Note a. Maximum score = 43

The ANOVA results indicated statistically significant differences between the groups: $F(2, 42) = 4.89$, $p = .012$. Tukey *a* post-hoc comparison test revealed significant differences between S-MWE-SE and NS-MWE groups (p = .021), and S-MWE and NS-MWE groups (p = .037). No significant differences between S-MWE-SE and S-MWE groups were found at p = .05 level.

Table **4.16**. Means and Standard Deviations for Time on Task

		Time on Task	
Group	n	M[a]	SD
Segmented Multimedia Worked Examples and Self-Explanation Prompts (S-MWE-SE)	12	117.08	62.97
Segmented Multimedia Worked Examples (S-MWE)	17	86.71	37.32
Non-Segmented Multimedia Worked Examples (NS-MWE)	16	69.00	35.85
Total	45	90.93	45.38

Note a. Time in minutes

The ANOVA results indicated statistically significant differences between the groups: The results were statistically significant: $F(2, 42) = 3.93$, $p = .027$. Tukey *a* post-hoc comparison test revealed significant differences between S-MWE-SE and NS-MWE groups ($p = .021$). No significant differences between S-MWE-SE and S-MWE groups or S-MWE and NS-MWE groups were found at $p = .05$ level. However, as indicated by the results of the box plots, there was an outlier found for S-MWE-SE group: with the outlier removed the descriptive statistics for this group were $M = 102.27$, $SD = 38.29$. With the outlier removed, ANOVA results were insignificant: F (2, 41) = 2.70, $p = .079$.

Since some of collaborating participants demonstrated somewhat similar self-explanation patterns within their collaborative groups, interrelationships among self-explanation dimensions and performance on the tests were also re-examined for the non-collaborating group. Table **4.17** provides the self-explanation component correlations for the non-collaborating group.

Table **4.17.** Correlations between Self-Explanation Dimensions and Test Performance (n = 12)

	Procedure-based	Principle-based	Quality	Correctness	Conceptual Knowledge	Problem Solving
Procedure-based	–	.006	.052	.244	.175	-.109
Principle-based		–	.887**	.600*	.171	.666*
Quality			–	.620*	.166	.505
Correctness				–	.529	.208
Conceptual Knowledge					–	.485
Problem Solving						–

Note. ** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).

To explore the relationships between the self-explanation dimensions and learning outcomes that did not achieve significance through linear correlation analysis, scatter plots were produced with each of the dimensions as independent variables and conceptual knowledge and problem solving performance measures as dependent variables, and the line was fitted to determine the nature of the relationship. Figures **4.7-4.8** demonstrate the relationships between the conceptual knowledge and problem solving test and those self-explanation variables where it was possible to determine the nature of the relationship. As expected from the correlation analysis results, there was a linear relationship between correctness of self-explanations and conceptual knowledge (Figure **4.7**), as well as quality of self-explanations and problem solving (Figure **4.8**). No meaningful non-linear relationships were found from the scatter plots for the non-collaborating group.

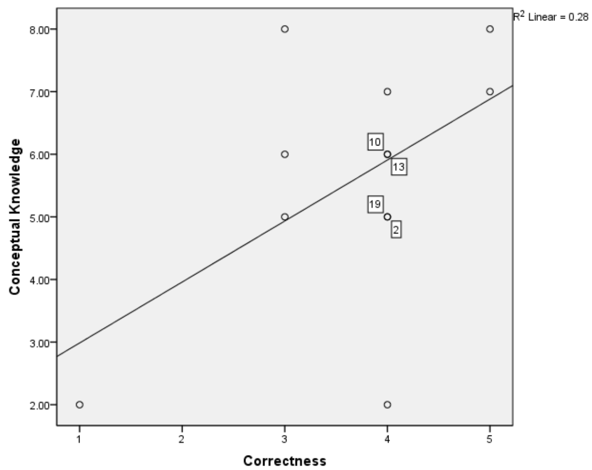


Figure 4.7. Correctness vs. Conceptual Knowledge

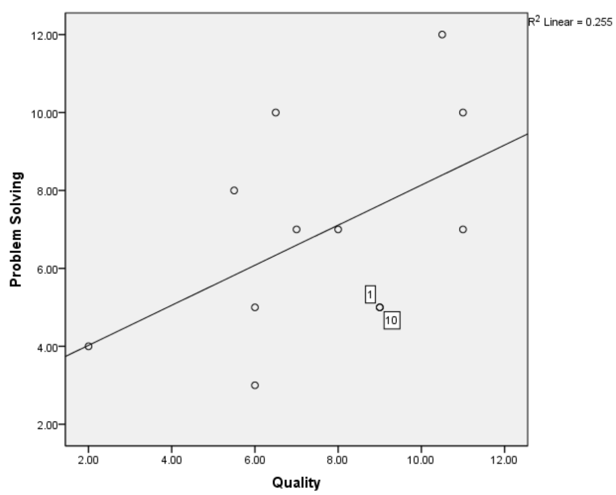


Figure 4.8. Quality vs. Problem-Solving Performance

A multiple regression analysis was conducted to determine whether any of the self-explanation dimensions (procedure-based explaining, principle-based explaining, quality, and correctness) were significant predictors of the performance on the tests. The stepwise regression method was selected because the correlation results suggested possible multicollinearity among several of the predictor variables (Kutner et al., 2005). No model could be produced for the conceptual knowledge measure. A model for the problem solving measure is presented in Table **4.18**. Principle-based self-explaining significantly predicted problem solving scores: *F Change* (1, 10) = 7.98, *Sig. F Change* = .018.

Table **4.18.** Multiple Regression Results for Interrelationship Between SE Dimensions and Problem-Solving Performance (n = 12)

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.666[ab]	.444	.388	2.121	.444	7.983	1	10	.018

Note a. Predictors: (Constant), Principle
Note b. Excluded: Procedure, Quality, Correctness

Comparing the Results of Collaborating vs. Non-Collaborating Participants

To determine if there were any difference in performance between those who collaborated ad those who did not, descriptive statistical tests and ANOVA were conducted for 45 non-collaborating participants and 17 collaborating participants for the conceptual knowledge test, problem-solving test, the treatment, and time on the treatment. Tables **4.19-4.22** provide the descriptive results:

Table 4.19. Means and Standard Deviations for Performance on Conceptual Knowledge Test

		Conceptual Knowledge	
Group	n	M[a]	SD
Non-Collaborating Group	45	5.56	2.28
Collaborating Group	17	8.18	2.35
Total	62	6.87	2.32

Note a. Maximum score = 10

Student t-test for independent samples revealed the significant differences between the groups: $t(60) = -4.00$, $p = .000$ with the collaborating group performing significantly better than the non-collaborating group.

Table 4.20. Means and Standard Deviations for Performance on Problem Solving Test

		Problem Solving	
Group	n	M[a]	SD
Non-Collaborating Group	45	7.80	2.54
Collaborating Group	17	7.65	2.45
Total	62	7.72	2.49

Note a. Maximum score = 13

Student t-test for independent samples did not reveal any significant differences between the groups: $t(60) = .21$, $p = .83$.

Table 4.21. Means and Standard Deviations for Problem Solving Performance in Treatment

		In-Treatment Performance	
Group	n	M[a]	SD
Non-Collaborating Group	45	27.71	5.18
Collaborating Group	17	29.59	3.48
Total	62	28.65	4.33

Note a. Maximum score = 43

Student t-test for independent samples did not reveal any significant differences between the groups: $t(60) = . -1.38$, $p = .17$.

Table 4.22. Means and Standard Deviations for Time on Task

		Time on Task	
Group	n	M[a]	SD
Non-Collaborating Group	45	88.51	47.94
Collaborating Group	17	76.82	28.73
Total	62	82.67	38.34

Note a. Time in Minutes

Student t-test for independent samples also did not reveal any significant differences between the groups: $t(48) = 1.17$, $p = .25$. There was an outlier in the non-collaborating group: with the outlier removed the descriptive statistics for that group were $M = 84.16$, $SD = 38.47$.

Patterning Multimedia Usage Behaviors

To determine patterns of multimedia usage and examine whether different patterns led to different performance outcomes, k-means clustering was used, where feature dimensions were the number of times the complete multimedia presentation and each multimedia segment were accessed.

Initially, there were two groups whose interactions with the multimedia example were determined by the study design: the first group took either the S-MWE-SE or S-MWE treatments and had access to both the complete multimedia presentation and multimedia segments, and the second group took the NS-MWE treatment which only had access to the complete multimedia presentation. Since group one participants could control which segments to access and how many times, it was anticipated that the patterns of usage in the segmented multimedia group would vary.

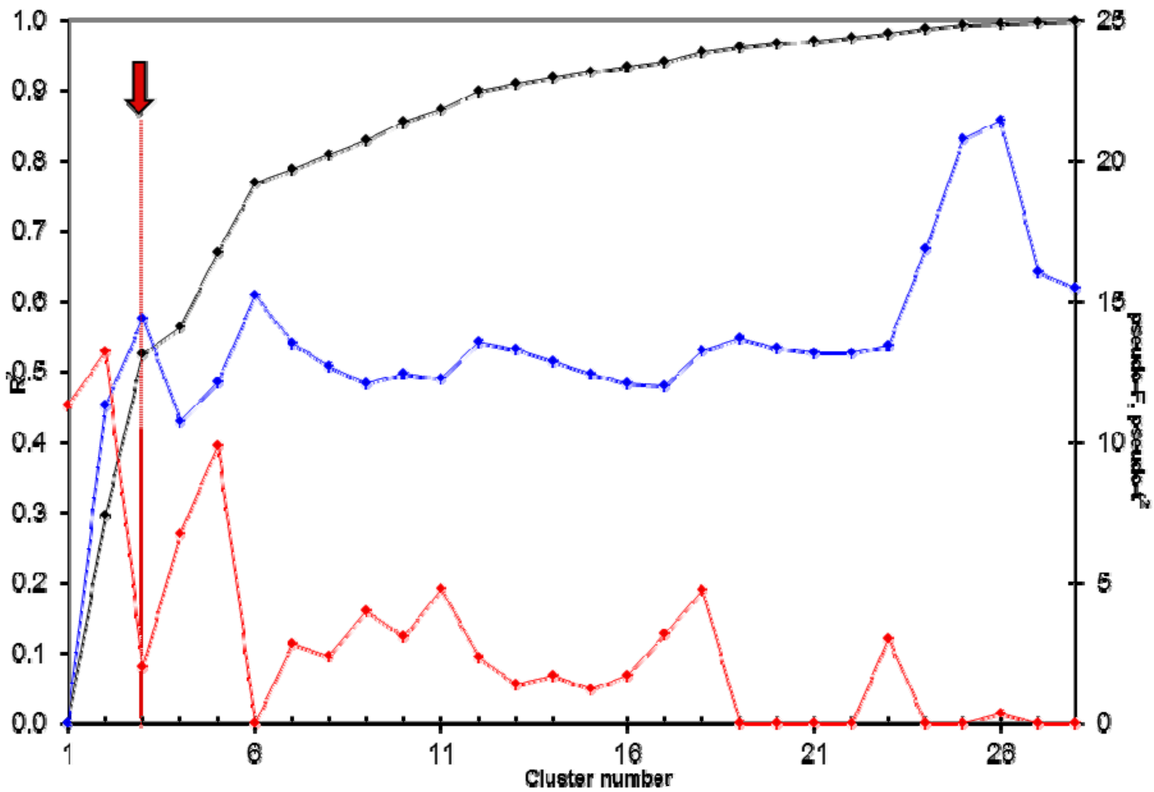
As the logs were being examined, it was discovered that there was a group of participants who were collaborating on the treatment. After initial exploration, it appeared that collaborating participants in the segmented multimedia groups accessed the multimedia segments significantly less: $t(40) = 3.76$, $p = .001$ (see Table **4.23** for descriptive results). This result, in combination with the evidence of collaboration, suggests that these participants' interactions with the treatments were different from the rest of the segmented multimedia group. Therefore, it was determined that the collaborating participants belonged to the group of their own, and they were excluded from the cluster analysis.

Table **4.23.** Means and Standard Deviations for Multimedia Segment Access

		Segment Access	
Group	n	M[a]	SD
Non-Collaborating Group	29	8.38	4.59
Collaborating Group	13	2.54	4.74
Total	42	5.46	4.66

Note a. Maximum number is unlimited

To determine patterns of interactions, the two-stage procedure described in Chapter 4 was used with 29 feature vectors for 29 non-collaborating participants from the segmented multimedia groups. The average linkage procedure was conducted for 7 feature dimensions that corresponded to 7 multimedia segments. The results in Figure **4.9** revealed 3 distinct clusters, as indicated by the red arrow.



Note. R^2 – black; *pseudo-F* – blue; *pseudo-*t^2 – red

Figure **4.9.** Number of Clusters Identified by Average Linkage Procedure

Then *k*-means clustering was performed for $k = 3$ with the initial cluster centers identified by the average linkage procedure. Then the final cluster centers were examined to identify what characteristics each cluster possessed. As a result, the clusters could be classified as follows: "most frequent viewers" (MFV) – 4 participants; "moderate viewers" (MV) – 15 participants, and "least frequent viewers" (LFV) – 10 participants. The final cluster centers are displayed in Table **4.24**.

Table **4.24.** Final Cluster Centers – Viewing Patterns

	Final Cluster Centers [a]		
	MFV	MV	LFV
Segment 2	5.00	2.73	1.0
Segment 3	2.50	1.07	.40
Segment 4	4.75	2.93	.90
Segment 6	.75	.53	.30
Segment 7	1.5	.93	.30
Segment 8	1.25	.87	.40
Segment 9	.25	.53	.10

Note a. Variable coordinates / average number of viewings per segment selected by participants

These 3 clusters were analyzed using ANOVA with the conceptual knowledge and problem solving tests as dependent variables to determine which patterns of interaction behaviors were more effective or less effective. Also, descriptive analysis and ANOVA were examined for possible differences in treatment scores and time on task. Tables **4.25 – 4.28** demonstrate the descriptive results for the three clusters.

Table 4.25. Means and Standard Deviations for Performance on Conceptual Knowledge Test

		Conceptual Knowledge	
Group	n	M[a]	SD
Most Frequent Viewers (MFV)	4	3.50	1.73
Moderate Viewers (MV)	15	5.67	2.26
Least Frequent Viewers (LFV)	10	6.10	2.08
Total	29	5.52	2.23

Note a. Maximum score = 10

The ANOVA results indicated no statistically significant differences between any of the groups: $F(2, 26) = 2.18$, $p = .13$.

Table 4.26. Means and Standard Deviations for Performance on Problem Solving Test

		Problem Solving	
Group	n	M[a]	SD
Most Frequent Viewers (MFV)	4	4.00	1.73
Moderate Viewers (MV)	15	8.33	2.16
Least Frequent Viewers (LFV)	10	7.70	2.58
Total	29	7.52	2.59

Note a. Maximum score = 13

The ANOVA results revealed statistically significant differences between the groups: $F(2, 26) = 6.10$, $p = .007$. Tukey *a* post-hoc comparison test revealed significant differences between the "most frequent viewers" and "moderate viewers" ($p = .005$), and "most frequent viewers" and "least frequent viewers" (p = .024). No significant differences between moderate access and least frequent access groups were found at p = .05 level.

Table 4.27. Means and Standard Deviations for Problem Solving Performance in Treatment

		In-Treatment Performance	
Group	n	M[a]	SD
Most Frequent Viewers (MFV)	4	29.00	3.63
Moderate Viewers (MV)	15	29.33	4.55
Least Frequent Viewers (LFV)	10	29.50	4.55
Total	29	29.34	4.10

Note a. Maximum score = 43

The ANOVA results indicated no statistically significant differences between any of the groups: $F(2, 26) = 0.02$, $p = .98$.

Table 4.28. Means and Standard Deviations for Time on Task

		Time on Task	
Group	n	M[a]	SD
Most Frequent Viewers (MFV)	4	119.75	50.22
Moderate Viewers (MV)	15	103.20	57.08
Least Frequent Viewers (LFV)	10	85.20	41.28
Total	29	99.28	50.85

Note a. Time in minutes

The ANOVA results indicated no statistically significant differences between any of the groups: $F(2, 26) = .74$, $p = .49$. There was an outlier in the MV group: with the outlier removed the descriptive statistics were *M* = 90.57 and *SD* = 30.53; ANOVA results with the outlier removed remained insignificant.

Summary of Findings

Findings from this chapter are summarized in Table **4.29**.

Table **4.29**. Summary of Findings

	All Participants	Non-Collaborating Participants	Frequency of Multimedia Access
RQ1: Do different types of multimedia worked examples influence acquisition of conceptual knowledge?	No, all three groups performed equally: $F(2, 59) = .00$, $p = .99$.	No, all three groups performed equally: $F(2, 42) = .02$, $p = .98$.	Multimedia Segment Access frequency did not affect conceptual knowledge performance: $F(2, 26) = 2.18$, $p = .13$.
		Collaborating participants scored significantly higher than non-collaborating participants: $t(60) = -4.00$, $p = .000$.	
RQ2: Do different types of multimedia worked examples influence problem solving performance?	No, all three groups performed equally: $F(2, 59) = 1.04$, $p = .36$.	No, all three groups performed equally: $F(2, 42) = 1.08$, $p = .35$.	Students from the segmented multimedia groups who accessed the multimedia segments the most frequently scored significantly lower: $F(2, 26) = 6.10$, $p = .007$. Tukey *a*: MFV and MV ($p = .005$); MFV and LFV ($p = .024$).

		No, both non-collaborating and collaborating groups performed equally: *t*(60) = .21, *p* = .83.	
RQ3: Is there an interrelationship between the type and quality of self-explanations, acquisition of conceptual knowledge, and problem solving performance?	Yes, a significant linear relationship between quality of self-explanations and problem-solving performance: Pearson's r = .461 R^2 = .17, *F change* (1, 17) = 4.60, *Sig. F change* = .047	Yes, a significant linear relationship between principle-based explaining and problem-solving performance for non-collaborating participants: Pearson's r = .666; R^2 = .39, *F change* (1, 10) = 7.98, *Sig. F change* = .018	N/A
	Yes, a non-significant linear relationship between principle-based explaining and problem-solving performance: Pearson's *r* = .436.	Yes, a non-significant linear relationship between quality of self-explanations and problem-solving performance for non-collaborating participants: Pearson's *r* = .505.	N/A
	Yes, a non-significant linear relationship between correctness of self-explanations and acquisition of conceptual knowledge: Pearson's *r* = .440.	Yes, a non-significant linear relationship between correctness of self-explanations and acquisition of conceptual knowledge for non-collaborating participants: Pearson's *r* = .529.	
	Yes, a possible non-linear relationship between procedure-based explaining and acquisition of conceptual knowledge		
	Yes, a possible non-linear relationship between principle-based explaining and		N/A

	acquisition of conceptual knowledge		
Other Findings			
In-Treatment Performance	S-MWE-SE group significantly outperforming the non-segmented group: *F*(2, 59) = 5.22, *p* = .008. Tukey *a*: *p* = .006	Both segmented groups significantly outperformed the non-segmented group: *F*(2, 42) = 4.89, *p* = .012. Tukey *a*: S-MWE-SE and NS-MWE groups (*p* = .021), and S-MWE and NS-MWE groups (*p* = .037). Non-collaborating and collaborating groups performed equally: *t*(60) = . -1.38, *p* = .17.	Multimedia Segment Access frequency did not affect in-treatment performance: *F*(2, 26) = 0.02, *p* = .98.
Time on Task	All three groups performed equally: *F* (2, 59) = 1.86, *p* = .17.	S-MWE-SE group spent significantly more time on task than the NS-MWE group: *F*(2, 42) = 3.93, *p* = .027. Tukey *a*: *p* = .021. Both collaborating and non-collaborating groups performed equally: *t*(48) = 1.17, *p* = .25.	Multimedia Segment Access frequency did not affect time on task: *F*(2, 26) = .74, *p* = .49.
Multimedia Access	No significant differences in complete example, unique segments, or total segments access among different MWE groups.	Collaborating participants accessed multimedia segments significantly less frequently: *t*(40) = 3.76, *p* = .001	There were three different multimedia access patterns among non-collaborating participants who received segmented MWE treatments.

Chapter 5

DISCUSSION AND IMPLICATIONS

Introduction

The study investigated the effects of non-segmented multimedia worked examples, segmented multimedia worked examples, and segmented multimedia worked examples enhanced with self-explanation prompts on acquisition of conceptual knowledge and problem solving performance in an undergraduate engineering course. In addition, the study examined learner-generated self-explanations for possible interrelationships between the type and quality of self-explanations and learning outcomes.

A total of 62 engineering students from a 300-level civil engineering course participated in the study as a part of their course. After the initial introduction of the concepts by the instructor in class, they completed the treatment and the conceptual knowledge test online as a part of their homework assignment, and completed a problem-solving test in class as a part of a course quiz. Participants' homework logs were retrieved to examine the generated self-explanations, and determine time on task and amount of access to the multimedia files.

A one-way ANOVA test did not reveal any significant differences between any of the groups on either outcome measure. Time on task also did not reveal any significant differences between the groups. However, in-treatment problem solving performance scores showed a significant difference between the S-MWE-SE and the NS-MWE groups with S-MWE-SE group performing significantly better.

There were several positive linear relationships found between principle-based type of explaining, quality of explaining, correctness of self-explanations and problem solving performance, as well as between correctness of self-explanations and acquisition of conceptual knowledge. Multiple regression results revealed that quality of self-explanations was a significant predictor of performance on the problem-solving test for the S-MWE-WE group.

Multimedia access logs from groups that received segmented multimedia example treatments were examined for possible patterns using a two-step clustering procedure. As a result, three distinct groups were identified based on their amount of access to each of the 7 segments: the most frequent viewers, moderate viewers, and the least frequent viewers. One-way ANOVA revealed that the most frequent viewers performed significantly worse on the problem-solving test than the other two groups. The results followed a similar pattern on the conceptual knowledge test, but the differences did not reach the significance level. However, all three types of viewers received very similar scores on the in-treatment performance.

It was also discovered from the logs that several participants may have collaborated on the treatment. Both collaborating and non-collaborating groups were examined to determine whether any possible differences in performance existed. There was a significant difference in performance on the conceptual knowledge test with the collaborating group performing significantly better. Multiple regression results revealed that principle-based explaining was a significant predictor of performance on the problem-solving test for non-collaborating participants in the S-MWE-WE group.

Discussion of findings

Effects of Multimedia Worked Examples (MWE) on learning outcomes

Worked examples have been established as a powerful strategy to teach problem solving at the initial stages of expertise development (Sweller & Cooper, 1985; Cooper & Sweller, 1987; Tarmizi & Sweller, 1988; Ward & Sweller, 1990). According to these authors, worked examples facilitated schema acquisition and rule automation by eliminating the need for means-ends analysis, and thus reducing extraneous cognitive processing. There was also evidence that as the learners' level of expertise increased, they started to rely less on worked examples (Kalyuga et al., 2001) and more on already developed schemas (Sweller et al., 1998). These empirical findings were also consistent with Anderson's (1995) 4-stage expertise development theory, which was based on the assumption that at the first stage learners solve problems by analogy, at the second stage they develop a schema, at the third stage schemas became more proceduralized, and at the fourth stage schemas became automatic, and the increased level of expertise and experience allowed for operating from a large pool of different types of examples using a higher level of analogical reasoning. The assumptions regarding the mechanisms of initial development of schemas were also supported by the analogical reasoning theory (Holyoak, 2005). In the current study it was found that segmenting multimedia worked examples did not result in higher learning outcomes for either acquisition of conceptual knowledge or problem solving performance. However, further exploration of learners' interactions with the treatments indicated that the learners exposed to segmented multimedia worked examples (S-MWE) performed significantly better on the in-treatment problem-solving test. The analysis of multimedia access patterns also revealed that the learners who were accessing the multimedia segments the most frequently, scored significantly lower on the problem-solving than those who were accessing the segments less frequently. The lack of the expected effect might be attributed to the following reasons.

1. Completing the Expertise Development Process

Examination of learners' problem solving performance within the treatments indicated significantly higher scores for one of the segmented multimedia groups (S-MWE-SE) compared to the non-segmented group (NS-MWE) for the full sample and for both segmented multimedia groups (S-MWE-SE and S-MWE) compared to the non-segmented group (NS-MWE) for non-collaborating participants. This finding suggests that the treatments were effectively supporting the problem-solving process, but these effects did not extend to the testing situation. A possible explanation of this finding might be that MWE was successful at the first stages of expertise development: solving problems by analogy and developing a schema, according to Anderson (Anderson, 1995; Anderson et al., 1997), or initial schema acquisition, according to Sweller (Sweller & Cooper, 1985; Cooper & Sweller, 1987; Sweller et al., 1998). However, it may have failed at the next stage of expertise development: schema automation and proceduralization, possibly due to insufficient opportunities for practice of the new skill within the learning environment. At the same time, both conceptual knowledge and problem solving tests were assessing performance at the fourth stage, operating automatic schemas (Anderson, 1995), which the learners may not have yet reached. These possible effects of multimedia worked examples are reflected in Figure **5.1**.

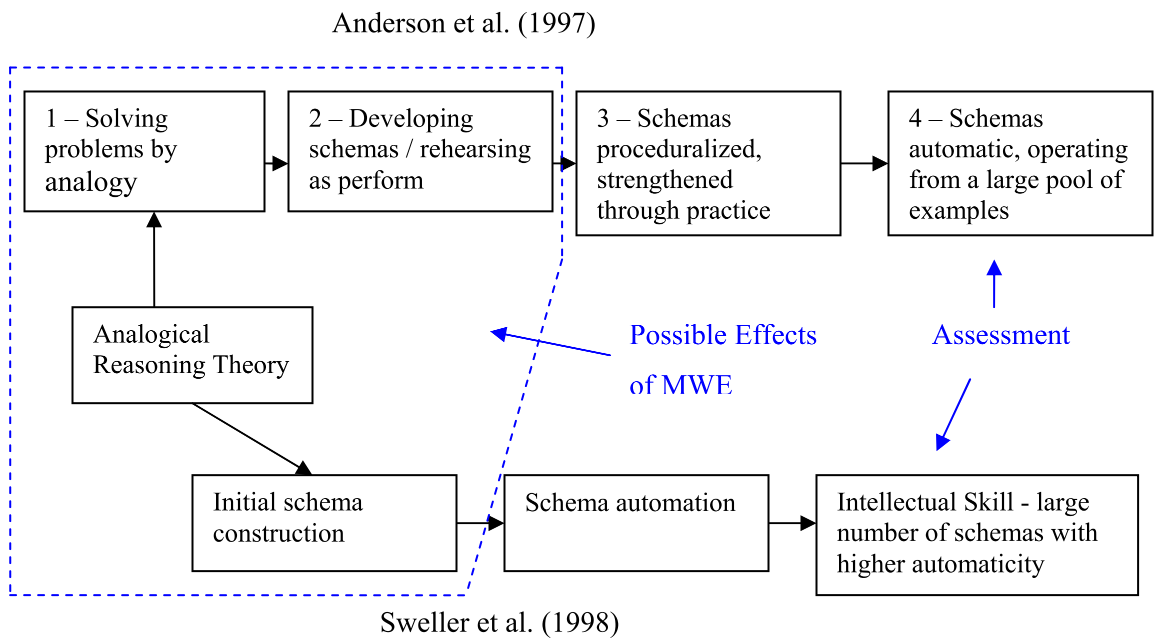


Figure **5.1.** Effects of MWE on Expertise Development Process

2. Scaffolding: Implications of Learner-Controlled Fading

Supporting problem-solving process with multimedia worked examples fits the definition of scaffolding. Scaffolding is "a process that enables a child or novice to solve a problem, carry out a task or achieve a goal which would be beyond his unassisted efforts" (Wood, 1976, p. 90). Pea (2004) brought the attention of the community to the original definition of scaffolding, which has its roots in Vygotsky's (1978) theory of zone of proximal development (ZPD). According to Pea, scaffolding is an adaptive process where there is a diagnosis of the learner's current proficiency, based on which adaptive level of support is provided by instructor or peer, with scaffolding activities "fading" as the level of proficiency increases. With this definition, Pea emphasized the difference between scaffolding and distributed intelligence, or distributed cognition, which does not include fading and is designed primarily to achieve activities that would be error prone, challenging, or impossible otherwise. Ability to eventually perform the task without supports is what ultimately constitutes the success of scaffolding as an instructional tool.

In this study, there was no fading built in the design, but learners could use their own judgment whether and to what extent to use the complete multimedia presentation and the supporting multimedia segments at each stage of the problem-solving process. However, an individual learner's ability to effectively fade the supports on their own remains an open question.

Closer examination of multimedia segments access for the two segmented MWE groups revealed that extensive access to the multimedia segments may have helped achieve adequate performance in the treatment materials, but did not result in effective performance on the outcome measures. However, those learners who relied heavily on instructional supports (Most Frequent Viewers) during the learning process performed significantly worse on the problem-solving test. One of the possible explanations might be that these learners did not attempt to solve the problems without assistance (Atkinson et al, 2003), and therefore did not develop the skill of performing the task independently relying on their new level of understanding. On the other hand, participants in the non-segmented group (NS-MWE) scored as well as and even 13% to 17% higher (depending on the sample) than the segmented with self-explanation group (S-MWE-SE) on the problem-solving test. One of the possible explanations might be that since they did not have as many multimedia supports as the other two groups, they may have learned how to perform without scaffolding.

3. Complexity of the Concepts and Skills

The conceptual knowledge-base and problem-solving skills acquisition of which this intervention was trying to facilitate are very complex and require building extensive conceptual understanding and a large pool of procedural schemas. To be able to make correct predictions as to how the systems would behave, i.e. knowing how the changes in process variables, such as flowrates, temperature, and pressure, will affect the performance of the system students should not only be able to apply problem-solving procedures and possess solid knowledge in the domain, but should also be able to integrate knowledge and process (Streveler et al., 2008). It is possible that these skills are difficult to develop from a single problem-set, and implementation of several problem-sets throughout the semester is necessary to better understand their effects on students' learning. Acquisition of expertise within the domain is typically gained over a long period or time, and expecting a major cognitive change to be accomplished within several days may not be entirely realistic (Brown, 1992).

4. Non-Segmented Treatment Design

All three groups received instruction that was breaking down the problem solving process into steps. It is possible that providing a step-by-step solution structure, i.e. organizing the information for the learners, was a powerful strategy (Mayer, 1999) that might have had a positive learning effect on NS-MWE group, and thus contributed to their equal performance regardless of treatment.

Effects of MWE+SE on learning outcomes

Self-explaining allows learners to monitor their own understanding, which makes it a powerful learning strategy (Chi et al. 1989). Explaining the reasons for their actions helps learners engage in deeper analysis of learning materials. The previous research on self-explanations suggests that not only the fact of self-explaining but also the type and quality of generated self-explanations make a significant difference in students' learning, especially conceptual understanding (Renkl, 1997; Roy & Chi, 2005; Hilbert et al., 2008). In their recent study, Berthold & Renkl (2009) found that both rationale-based self-explanations and principle-based self-explanations had positive effect on conceptual knowledge. However, there was conflicting evidence whether self-explanations had positive effect on procedural knowledge and problem-solving (Renkl, 1997; Berthold & Renkl, 2009). Lack of positive effects in some of the cases was attributed to increased cognitive load which interfered with building procedural schemas (Gerjets et al., 2006). The findings of the current study confirm only some the findings of past research regarding the effect of self-explanations: while no significant positive effect of self-explanation prompts in S-MWE treatment on learning outcomes was found per se, as the results of multiple linear regression revealed, the type of self-explaining and quality of the written self-explanations were significantly positively related with performance on the problem-solving test. The lack of the expected overall effect might be attributed to the following reasons.

1. Decontextualization of Principles May Require More Exposure to Various Examples and Practice Problems

Lack of effect of self-explanations in S-MWE on acquisition of conceptual knowledge suggests that self-explanation prompts did not facilitate far transfer. Far transfer is a "mindful generation of an abstraction during learning and its later application to a new problem or situation from which basic elements are similarly abstracted" (Salomon & Perkins, 1989, p. 127). According to Salomon & Perkins (1989), mindful abstraction is a deliberate decontextualization of a principle, main idea, strategy or procedure (p. 126).

The qualitative analysis of the content of generated self-explanations indicated that all of the participants engaged in self-explaining to some extent. Out of 95 responses, only 17 were coded as statements (18%), and the rest 78 (82%) were coded as self-explanations of different types and quality. At the same time, all 19 participants engaged in self-explaining for at least 2 out of 5 self-explanation prompts, with 15 participants (74%) engaging in self-explaining for at least 4 self-explanation prompts. Such high numbers might be attributed to the following factors: 1) self-explanation prompts in this study were conceptual and content-specific and addressed very specific concepts, which was considered a more effective strategy in the prior research (Berthold et al., 2004); 2) participants had an opportunity to model their responses from the instructor's explanations in the multimedia segments; and 3) participants were third-year engineering students who already had some experience with engineering mode of thinking that is based on conceptually-supported judgment (Streveler et al, 2008). This qualitative evidence suggests that most learners did engage at least to some extent in "mindful generation of an abstraction during learning" (Salomon & Perkins, 1989, p. 127), however it did not lead to application of these abstractions to the conceptual problems in the conceptual knowledge test. A possible explanation might be that students did not have enough opportunities to strengthen their newly generated abstractions by being exposed to either more worked examples, example-problem pairs, or at least practice problems with varying surface structure during the learning stage (Paas & Van Merrienboer, 1994b). In this study, the problems had varying surface structure, but there was only one problem per type. This lack of practice opportunities might have resulted in incomplete decontextualization of principles, i.e. insufficient understanding of the principles independent of a concrete example problem structure, which might have made far transfer difficult to achieve.

2. Principle-Based Explaining and Higher Quality of Self-Explaining are Associated with Better Problem Solving

The analysis of type and quality of written self-explanations revealed that 1) there was a significant positive relationship between the quality of self-explanations and performance on the problem-solving test for the complete MWE-SE group, and there is a significant positive relationship between the principle-based nature of self-explaining and performance on the problem-solving test for non-collaborating participants in the MWE-SE group; 2) there was a significant inter-relationship between principle-based type of explaining and the quality of self-explanations; 3) there was not enough evidence in support of the assumption that the type and quality of written self-explanations would be significantly positively correlated with the conceptual knowledge scores. These findings contradict some of the earlier research findings that suggest that improvement in conceptual understanding is primarily related to higher quality of self-explanations (Hilbert et al., 2008; Berthold & Renkl, 2009), and that self-explanations may even be detrimental for problem-solving performance (Berthold & Renkl, 2009). However they are consistent with the finding by Renkl (1997) that successful learners employed more principle-based explanations, more explication of operator-goal combinations, and more anticipative reasoning. A possible explanation for that finding may be that while complete decontextualization of principles did not occur, those learners who engaged in higher levels of self-explaining were able to apply their generated abstractions of the principles and procedures to solve near transfer problems.

3. Wording and Placement of Self-Explanation Prompts Make a Difference

Closer examination of written self-explanations revealed that not all five self-explanation prompts (SEP) triggered the same pattern of responses (see Table **4.3** and Figures **4.5** and **4.6** in Chapter 4). The primary purpose of SEPs was to solicit principle-based explaining to facilitate conceptual understanding (Renkl, 1997). However, as can be seen on Figures 4.4 and 4.5, SEP2, SEP3, SEP4 produced more principle-based than procedure-based explaining, SEP1 and SEP5 produced more procedure-based explaining than principle-based explaining, and at the same time, SEP4 produced the fewest number of self-explanations in general. During the design phase, SEPs like “What principle is involved in solving for this step?” were avoided not to trigger single-word responses; and questions “Explain why…” and “Justify your selection of the answer…” were used instead. While forcing learners to explain their decisions produced a high overall number of explanations (82%), it did not necessarily promote principle-based explaining for all SEPs.

Findings from the wording and placement analysis of the five self-explanation prompts (SEP), as well as possible explanations and recommendations are summarized in Table **5.1**.

Table 5.1. Findings and Recommendations for Self-Explanation Prompt Design

	More Principle-based explaining results		More Procedure-based explaining results		None
	SEP 2	SEP 3	SEP 1	SEP 5	SEP4
Current wording	Short and content-specific	Short and content-specific	Short and content-specific	Short and content-specific	Short and content-specific
Current placement	Principle is observable and easier to abstract	Principle is less observable but relatively easy to abstract from comparison	Procedure is salient in the demonstrated idea; principle is more difficult to abstract	Procedure is salient in the demonstrated idea; principle is more difficult to abstract	Procedure is salient in the demonstrated idea; principle is relatively easy to abstract
Principle	Application				
Differences in pressure depending on the measurement point	concrete and real life; context is relatively simple			context is more complex; easier to explain with a procedure	
Differences between a manometer and a pitot tube		involves comparing two systems without identifying them			too much conceptual information in the problem statement and procedural information in the analytical question

Selecting correct measurement points			"Selection" is heavily procedural and more obvious		
Recommendation					
Wording of prompt	Keep short but content-specific	Keep short but content-specific	Expand: ask to explain which principle is involved and why	Expand: ask to explain which principle is involved and why	N/A
Placement of prompt	Keep	Keep	Keep	Keep with an expanded prompt or change	Change or provide less conceptual and procedural information

From the conceptual standpoint, SEP2 and SEP5 both addressed the same principle (differences in pressure depending on the measurement point), and SEP3 and SEP4 addressed another principle (differences between a manometer and a pitot tube). SEP1 addressed a somewhat different principle (selecting correct measurement points). SEP1 essentially asked to explain a procedure, which could be explained either from the procedure or principle standpoint. As shown in Figure 4.3, procedure-based explaining for this SEP prevailed with a 33% difference between the scores. SEP2 and SEP5 were addressing the same principle in two somewhat different situations, and produced two very different patterns of responses: the first one was heavily principle-based with 67% difference between the two scores, and the second one was almost as heavily procedure-based with 59% difference between the scores. A possible explanation for that difference in pattern might be that in case of SEP2 the principle was more easily observed and less abstract (water coming out of the jet experiences atmospheric pressure which is virtually the same at both measurement points), while in case of SEP5 differences in pressure were measured inside the pipe and depended on changing of a number of parameters, which were more easily explained procedurally. SEP3 and SEP4 also addressed the same principle, but while SEP3 produced high amounts of both principle-based and procedure based explaining, with all 19 participants explaining rather than stating, SEP4 resulted in a low amount of explaining with 47% of statements (see Table **4.3**). A possible explanation for that is that SEP3 asked learners to compare two systems that differed in one parameter that was also easily observable, and the systems were not identified as a manometer and a pitot tube in the problem statement; at the same time, in case of SEP5, 1) there may have been too much conceptual information given in the problem statement; and 2) the numeric question preceding the SEP was already very detailed, and it may have been difficult to explain beyond the numeric response.

These findings suggest that there are certain characteristics of the self-explanation prompt design that make them more successful at eliciting higher amounts of self-explaining, higher quality of self-explaining, and more principle-based explaining. One of these characteristics might be placement: some steps in the problem-solving process might present better opportunities for placing self-explanation prompts: usually these are the steps where the principle is easier to abstract and does not compete for the explanation with a procedure, because in that case a procedure-based explanation will be the obvious first-choice. Another characteristic might be wording: while it may be sufficient to use a shorter and more general content prompt for a well-placed self-explanation, there may be parts in the problem-solving process where the principle is more difficult to extract, but which need to be supported by a self-explanation prompt. In that situation, wording of the prompt should be extended to ask to explain which principle is involved and why.

5. Interference of Multimedia Self-Examples with the Self-Explaining Process

Instructional explanations from the MWE segments might have interfered with the self-explaining process because the availability of those instructional explanations might have reduced the self-explanation effort of the learners, thus, negatively affecting the learning outcomes (Schworm & Renkl; 2006). Another possible negative effect of the combination of these two kinds of learning supports might be due to them increasing cognitive processing demands on the learners' working memory (Kalyuga et al, 2000; Sweller, 2005). However, high in-treatment scores for the S-MWE-SE group suggest that these explanations may not be completely exhaustive, and this issue needs further exploration.

Instrument Design Findings

Conceptual knowledge test: testing students online is a convenient option, but there were two main problems associated with administering tests outside of classroom: 1) a possibility that students might use books or other outside resources to enhance their scores; and 2) a possibility that two or more students might collaborate on the test. Several measures were taken to minimize the impact of the first issue: first of all, the conceptual problems were structured in a way that would require participants to use their engineering judgment, i.e. to rely more on their deeper understanding of the concepts rather than on books (Streveler et al., 2004; Streveler et al., 2008); second, a 20-min time limit was set for the completion of the test. To minimize the possibility of collaboration, learners were explicitly asked to work on online-related activities individually. The results of the test and examination of the logs indicated that there was some evidence of "ceiling effect", i.e. there was a high number of maximum or near- maximum scores (Uttl, 2005), but this effect seemed to be more attributed to students collaborating on the test rather than other factors (see Figures **5.2** to **5.4**).

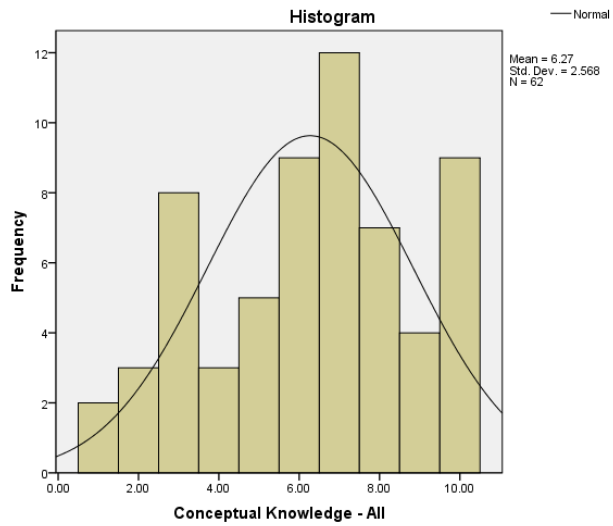


Figure 5.2. Conceptual Knowledge Scores Distribution for All Participants

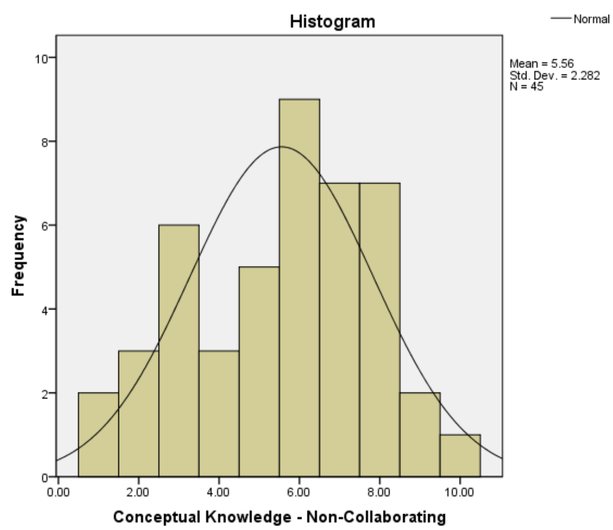


Figure 5.3. Conceptual Knowledge Scores Distribution for Non-Collaborating Participants Only

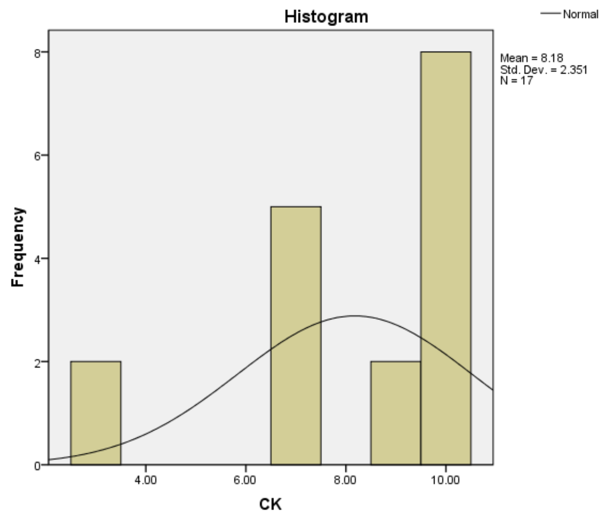


Figure 5.4. Conceptual Knowledge Scores Distribution for Collaborating Participants Only

Problem-solving test: the challenge of developing problem-solving assessments for an actual classroom environment was that the test could only be implemented during one 50-minute class session. According to the subject matter experts, this amount of time was not sufficient to solve three complex fluid mechanics problems. Instead, in the test used for this study, students were asked to partially solve 2 problems, and write out a complete solution for the third problem, with the third problem similar to the most difficult problem from the intervention. While the most difficult problem did encompass most of the concepts involved in the intervention, there were some issues associated with the complexity of it.

The analysis of the scores from the problem-solving test revealed that there was an area of procedural and conceptual misunderstanding that could have been more explicitly supported by either an additional multimedia segment or by a self-explanation prompt in the treatment. This issue was unanticipated by the researcher and the content experts, but learners' lack of understanding of this particular procedure impaired their ability to solve one of the problems in the problem-solving test: an assumption that a large number of students did not realize they needed to use one additional equation to solve the problem was supported by low item difficulty indices for items P3.3.1 and P3.3.2 (.1 and .2 respectively) which resulted in a low item difficulty index for item P3.5, which was .1 as well. A possible consequence of this issue might be that it was not the intervention that influenced some of the observed differences in the scores on the problem-solving test, but rather higher ability of individual learners.

Implications for Instructional Design

Based on the results of this study and the review of previous research, the following implications for instructional design can be drawn.

Implications for Intervention Design

First, segmented multimedia worked example (S-MWE) intervention did not result in higher problem solving performance or acquisition of conceptual knowledge when compared with a NS-MWE group. More problems may be needed in the problem-set to help learners not only to develop initial schemas but also to achieve schema automation. However, there is an important issue associated with this recommendation: it is the complexity of the problems and time that is necessary to solve them. Increasing the problem-set to 5-6 problems to allow enough practice opportunities is likely to significantly increase time on task. To alleviate this issue, several smaller problem-sets may need to be introduced over a longer period of time such as a month or a semester. To facilitate the transition from a learning situation to a testing situation, fading may need to be used to gradually remove supports starting either backwards (Atkinson et al., 2003; Renkl et al., 2004), or starting from the least difficult solution steps. This strategy may help direct attention for those students who need more guidance in their learning process, and also will help learners practice schemas with less and less support.

Second, enhancing multimedia worked examples with self-explanation prompts did not result in superior performance on the conceptual knowledge and problem-solving tests over MWE alone. Since the results demonstrated that principle-based explaining and higher quality of generated self-explanations were associated with better performance on the problem-solving test, self-explanation prompts should be content-specific and positioned at such places in the problems where they would solicit the most principle-based explaining from the learners. To facilitate acquisition of conceptual knowledge, more opportunities may need to be provided for learners to explain the same principle applied in various situations through either increasing the number of example-problem pairs or at least the number of practice problems.

Third, it may be important to pay attention to the design and placement of the self-explanation prompts. When a principle is easily observable, it may be enough to have a short self-explanation prompt, such as “Explain your answer”, “Justify your selection”, etc. But if the principle is less observable and more abstract, more guidance may be needed, such as including words “Explain from the principles standpoint” or “Justify using principles”. However, not all the analytical questions can be supported with a self-explanation equally well: even though most questions address a certain underlying principle, some calculation, comparison, and evaluation questions may not provide a rich background for principle-based self-explaining, and may require more extensive probing than the written self-explanation format allows for.

Implication for Assessment Design

Test design: In the test used for this study, students were asked to partially solve 2 problems, and write out a complete solution for the third problem, which was similar to the most difficult problem from the treatment. For the future, to access the learning outcomes more precisely, and also to reduce the number of extremely difficult items, students should be asked to write out complete solutions for two less difficult problems and partially solve the more difficult one.

Test administration: testing students online is a convenient option, but there are two main problems associated with administering tests outside of classroom: 1) a possibility that students might use books or other outside resources to enhance their scores; and 2) a possibility that two or more students might collaborate on the test. To avoid such issues, it may be better to either administer tests in the classroom / supervised lab, or use other assessment methods that are less sensitive to non-controlled environments.

Implications for Research Design

Content specialists and faculty have different needs than the researcher

The initial goal of the project was to support students' problem solving process with a segmented multimedia worked example while they were working on their homework. Eventually, self-explanation prompts were added, and then means to assess the outcomes beyond the homework grade had to be devised. While both the course faculty, the subject matter experts, and the researcher were working towards the same goal – improving classroom instruction, there were some compromises that had to be made to accommodate everybody's needs.

Experimental design: there are certain inconveniences and ethical issues associated with assigning students to different treatment groups. There were two convenience issues associated with implementation of an experimental design in the actual classroom: 1) designing three versions of the intervention in the course management system was cumbersome and required many resources of the instructional designer, the instructor, and the teaching assistants; 2) retrieving and consolidating homework results on each of the three problems from three groups was more time-consuming relative to the regular course practices. From an ethical standpoint, experimental design can potentially create significant differences in scores and thus put some of the learners at a disadvantage. To alleviate this issue, an initial solution was proposed that, though the study would be conducted only on the first problem-set, the experimental design would be applied all problem-sets used during the semester with alternating the treatments between the groups. However, after the study on the first problem-set was conducted, the instructor made the decision not to use the experimental design for the rest of the problem-sets, due to the convenience issues discussed above and also because more than two other problem-sets were implemented: there were a total of five problem-sets delivered during the semester. Instead, all multimedia worked examples were opened for review before the final exam.

Measuring learning outcomes: extensive testing, while necessary for the research purposes, is not viewed as very desirable by most faculty who do not want to add to the workload of their students, considering the overall amount of concepts and procedures that the students are being tested on during one quiz and during the course of a semester. A compromise had to be made on the number of items in both conceptual knowledge and problem solving tests.

Time and resource limitations

Since the research had to follow the timeline and the schedule of the actual class, several compromises had to be made:

1) The pilot study was conducted on a problem-set that was scheduled to be delivered towards the end of the semester. Since previous exposure to a multimedia worked example and self-explanation prompts would have made it difficult to discern the effects of the treatments, it was decided to conduct the main study on the newly developed problem-set that was scheduled to be delivered in the beginning of the semester. As the timeline for producing all the materials for the new problem-set was very short, there was not enough time to validate the instruments before the main study.

2) Though the findings from the previous research found example-problem pairs the most effective (Renkl, 2005), there was not enough time and resources to develop 3 different examples to support each of the 3 problems in the problem-set individually – instead, parts of the first example were completely supporting Problem 1 and partially supporting Problem 3, and the second example was not broken into steps, and was supporting Problem 2 and partially Problem 3.

Actual Classroom and Online Environment

The essence of the homework process is that it is independent work outside of classroom. It can be done anywhere, any place during a certain timeframe between the time when the homework is assigned and the time by which it has to be submitted. While the online environment allowed for tracking some of the learners' actions, there were a number of areas that were out of the researcher's control. First of all, it was impossible to determine how much of the time that the learners spent on the treatments was off task. Second, though the learners were explicitly asked to work on the treatments alone, there was some evidence in the logs that suggested that between one third and one fourth of them might have been collaborating on the treatment and one of the tests with at least one more person. Both distractions and collaboration are typical for homework assignments. While they cannot be controlled, some of these interactions can be tracked and accounted for in the research process. There were other factors typical for the actual classroom and out of the researcher's control, such as individual students possibly attending office hours, discussing the concepts from the homework (i.e. the treatment) with their classmates, reading the textbook, or solving additional problems before the test. Random assignment to treatments was used to help minimize the effects of these individual differences in learning strategies on the learning outcomes.

Also, since the research study was conducted as a part of a classroom intervention, the pool of participants was limited to the number of students enrolled in the course, which was around 80 students.

Limitations of the Study

The study was conducted as a part of a classroom research project, and all participants in the sample came from the same undergraduate engineering class within a large state university. The pool of participants was also limited by the size of the class. Since pursuing an engineering degree is required for enrollment in the class, the results can only be generalized to other civil engineering students or undergraduate students in similar engineering disciplines. Also, this study was conducted within the constraints of a specific classroom, and while this was a typical undergraduate mid-size course, there might have been some unique factors that contributed to the results of the study.

Recommendations for Future Research

Research Design

a. *Examining the effects of different types of MWE fading on acquisition of conceptual knowledge and problem-solving performance.* In past research, fading was found to be a powerful strategy to facilitate schema automation when learning from worked examples. Solving complex engineering problems might require non-linear fading of the MWE segments, as opposed to backward or forward fading. It is also worth exploring whether adding more problems to the treatment with or without fading will produce better learning outcomes.

b. *Using repeated measures design with several treatments throughout the semester to determine long-term effects.* Deep conceptual understanding and solid problem-solving skills usually require more time to develop than one week. Implementing several problem-sets during the semester should provide more insights on whether MWE and SE significantly affect long-term learning outcomes. Tracking self-explanations generated by the learners over an extended period of time might also provide researchers with more information about possible changes in the type and quality of self-explaining as learners' conceptual understanding and problem-solving skills develop over the course of semester.

c. *Examining the effects of self-explanation prompts and type/quality of self-explanations with segmented and non-segmented multimedia worked examples.* It is possible that use of MWE contributed to the lack of effect of self-explanations, because of either instructional explanations reducing learners' self-explanation effort or MWE and self-explanations competing over the learners' cognitive processing capacity. Using a 2 x 2 factorial design should help determine whether the combination of the two learning supports reduces their individual benefits.

d. *Adding a self-regulating strategies assessment.* An MSLQ could be added to examine whether individual cognitive strategies contribute to differences in learning outcomes for different types of MWE and influence the type of self-explaining and quality of self-explanations. Since self-explanation prompts aim at engaging more advanced cognitive strategies, such study might help identify the source of differences in performance on self-explanation treatments for individual learners.

e. *Further exploring the effects of collaboration on learning from multimedia worked examples.* Collaboration on a homework assignment is a typical classroom situation, sometimes even encouraged by the instructor. While examining the effects of collaboration was not the purpose of this study, several findings indicated that collaborative groups interacted with the treatments differently. Determining whether such collaboration is productive for learning from MWE might be an interesting direction for future research in this area.

Conclusions

The purpose of the study was to investigate the effects of segmented multimedia worked examples and self-explanations on acquisition of conceptual knowledge and problem solving performance in an actual undergraduate engineering classroom. The results showed that neither the segmented multimedia worked examples alone or enhanced with self-explanations significantly influenced conceptual understanding or problem solving. However, the results from the in-treatment problem-solving performance suggested that segmented multimedia worked examples and self-explanations did significantly influence the learning process, but not transfer.

These results suggest that segmenting alone and segmenting enhanced with self-explanation prompts may not be effective based on the provided amount of practice. More research is recommended to explore this explanation and the recommendation to increase the number of worked examples and practice problems to facilitate automation of schemas and decontextualization of principles.

The analysis of type and quality of self-explanations revealed a significant positive interrelationship between the type and quality of written self-explanations and problem-solving performance. Correctness of self-explanations was positively related with both acquisition of conceptual knowledge and problem solving performance. Based on the analysis of generated self-explanations and wording and placement of self-explanation prompts, it was possible to provide some recommendations for how to elicit higher quality of self-explaining and encourage principle-based explaining in a particular context.

Analysis of multimedia access patterns revealed that more frequent access to multimedia supports does not necessarily result in higher performance on the test. These results suggest that providing these supports without explicit and structured fading may not be effective. More research is recommended to explore this explanation and the recommendation to provide faded scaffolding to facilitate the transition to independent problem-solving for learners with possible self-regulation problems or lower problem-solving abilities.

Finally, the analysis of treatment logs suggested a high likelihood of collaborative work among some of the participants. Given significant differences in performance on the conceptual knowledge test, favoring the collaborating group, more research is recommended regarding the benefits of learners collaborating on problem-solving assignments for well-defined problems.

REFERENCES

Ainsworth. S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16,* 183 – 198.

Ainsworth, S. E., & Loizou, A. (2003). The effects of self-explaining when learning with text or diagrams. *Cognitive Science, 27(3),* 669–681.

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

Alexander, P.A., Schallert, D., & Hare (1991). Coming to Terms: How Researchers in Learning and Literacy Talk about Knowledge. *Review of Educational Research***,** *61,* pp. 315-343.

Alexander, P.A. (2003). The development of expertise: The journey from acclimation to proficiency. *Educational Researcher*, *32 (8)*, 10-14.

Amershi, S. & Conati, C. (2009). Combining unsupervised and supervised classification to build user models for exploratory learning environments. *Journal of Educational Data Mining, 1 (1),* 18-71.

Anderberg, M. R., (1973). *Cluster Analysis for Applications*, Academic Press.

Anderson, J. R. (1995). ACT: A simple theory of complex cognition. *American Psychologist, 51*, 355-365.

Anderson, J. R., Fincham, J. M., & Douglass, S. (1997). The role of examples and rules in the acquisition of a cognitive skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 932-945.

Anderson, L.W. & Krathwohl, D.R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessment: A revision of Bloom's taxonomy of educational objectives* (Abridged ed.). New York, NY: Longman.

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

Atkinson, R.K., Derry, S.J., Renkl, A. & Wortham, D. (2000). Learning from examples: Instructional principles from worked examples research. *Review of Educational Research, 70 (2),* 181 – 214.

Avramov, A. (2005). *Objective synoptic typing during the mixed-phase arctic cloud experiment.* Unpublished master's thesis, The Pennsylvania State University, University Park.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews, 4*, 829-839.

Baker, R.S.J.D. & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining, 1 (1),* 3-17.

Berthold, K., Nückles, M., & Renkl, A. (2004). Writing learning protocols: Prompts foster cognitive and metacognitive activities as well as learning outcomes. In P. Gerjets, J. Elen, R. Joiner, & P. Kirschner (Eds.), *Instructional design for effective and enjoyable computer-supported learning* (pp. 193-200). Tübingen: Knowledge Media Research Center.

Berthold, K. & Renkl, A. (2009). Instructional aids to support a conceptual understanding of multiple representations. *Journal of Educational Psychology, 101(1),* 70-87.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences, 2(2),* 141–178.

Catrambone, R., & Yuasa, M. (2006). Acquisition of procedures: The effects of example elaborations and active learning exercises. *Learning and Instruction, 16(2),* 139-153.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13(2)*, 145-182.

Chi, M. T. H., de Leeuw, N., Chiu, M. H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18,* 439–477.

Conati, C., & VanLehn, K. (2000). Toward computer-based support of meta-cognitive skills: A computational framework to coach self-explanation. *International Journal of Artificial Intelligence in Education, 11,* 398–415.

Cooper, G. & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology, 79*, 347-362.

Darabi, A.A., Nelson, D.W., & Paas, F. (2007). Learner involvement on a complex cognitive task: Application of a composite measure of performance and mental effort. *Journal of Research on Technology in Education, 40(10),* 39-48.

DeLeew, 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.

Ericsson, K A, & Lehmann, A C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology, 47,* 273-305.

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

Große, C. S., & Renkl, A. (2006). Effects of multiple solution methods in mathematics learning. *Learning and Instruction, 16(2),* 122-138.

Hilbert, T.S., Renkl, A., Kessler, S., & Reiss, K. (2008). Learning to prove in geometry: Learning from heuristic examples and how it can be supported. *Learning and Instruction, 18,* 54-65.

Holyoak, K.J. (2005). Analogy. In K.J. Holyoak & R.G. Morrison (Eds.) *The Cambridge Handbook of Thinking and Reasoning*. NY, NY: Cambridge University Press.

Huang, X. (2007). The effects of different types of worked examples on student learning and transfer of a problem-solving task. The Florida State University. Retrieved August 10, 2009, from http://ezaccess.libraries.psu.edu

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors, 40,* 1–17.

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

Kalyuga, S., Chandler, P., Tuovinen, J., & Sweller, J. (2001). When problem solving is superior to studying worked examples. *Journal of Educational Psychology, 93*(3), 579-588**.**

Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review, 19,* 239–264.

Kutner, M. H., Nachtsheim, C. J., Neter, J., & LI, W. (2005). Applied linear statistical models. (5th ed.). New York: McGraw-Hill.

Lovett, M.C. (2002). Problem Solving. In H. Pashler & D. Medin (Eds.) *Stevens' Handbook of Experimental Psychology Volume 2: Memory and Cognitive Processes, 3rd Ed*. NY, NY: John Wiley & Sons, Inc

Low, R. & Sweller, J. (2005). The modality principle in multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 147–158). Cambridge, England: Cambridge University Press.

Mahan, W.L. (2007). *Effect of prior knowledge and principle-based and procedure-based instructional strategies on students' transfer of principles.* The Pennsylvania State University. Retrieved August 14, 2008, from http://ezaccess.libraries.psu.edu

Markman, A.B. (2002). Knowledge representation. In H. Pashler & D. Medin (Eds.) *Stevens' Handbook of Experimental Psychology Volume 2: Memory and Cognitive Processes, 3rd Ed*. NY, NY: John Wiley & Sons, Inc.

Martin, J., Mitchell, J. & Newell, T. (2003). Development of a concept inventory for fluid mechanics. Proceedings of the Frontiers in Education conference, Boulder, CO, November, 2003

Mayer, R.E. (1999). Designing instruction for constructivist learning. In C.M. Reigeluth, *Instructional-design theories and models*. 2nd ed. Mahwah, NJ: Lawrence Erlbaum Associates.

Mayer, R. E. (2005a). Introduction to multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 1-16). Cambridge, England: Cambridge University Press.

Mayer, R. E. (2005b). Principles for managing essential processing in multimedia learning: Segmenting, pretraining, and modality principles. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 169–182). Cambridge, England: Cambridge University Press.

Mayer, R. E. (2005c). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 183–200). Cambridge, England: Cambridge University Press.

Mayer, R. E. (2005d). Principles of multimedia learning based on social cues: Personalization, voice, and image principles. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 201–212). Cambridge, England: Cambridge University Press.

Mayer, R.E. & Wittrock, M.C. (2006). Problem Solving. In P.A. Alexander & P.H. Winne (Eds.). *Handbook of Educational Psychology*, 2nd Edition. Lawrence Erlbaum.

Merceron, A. & Yacef, K. (2005). Educational data mining: A case study. *Proceedings of the 12th international Conference on Artificial Intelligence in Education.*

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81-97. Retrieved from: http://psychclassics.yorku.ca/Miller/

Milligan, G.W. (1980). An Examination of the Effect of Six Types of Error Perturbation on Fifteen Clustering Algorithms. *Psychometrika*, *45*, 325 -342.

Moreno, R. (2006). When worked examples don't work: is cognitive load theory at an impasse? *Learning and Instruction, 16,* 170-181.

Moreno, R. & Mayer, R.E. (2007). Interactive Multimodal Learning Environments. *Educational Psychology Review, 19,* 309–326.

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

Paas, F. G. W. C., & van Merrie¨nboer, J. J. G. (1994a). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review, 6,* 51–71.

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

Paivio, A. (1991). Dual coding theory: Retrospective and current status. *Canadian Journal of Psychology, 45,* 255 – 287

Pea, R.D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *The Journal of the Learning Sciences, 13(3),* 423–452.

Pintrich, P. R., De Groot, E. V. (1990). Motivational and self-regulated learning component of classroom academic performance. *Journal of Educational Psychology, 82*, 33-40.

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

Reisslein, J., Atkinson, R. K., Seeling, P., & Reisslein, M. (2006). Encountering the expertise reversal effect with a computer-based environment on electrical circuit analysis. *Learning and Instruction, 16(2),* 92-103.

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

Renkl, A. (2005). The worked-out-example principle in multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 229-246). Cambridge, UK: Cambridge University Press.

Renkl, A., Atkinson, R. K., & Große, C. S. (2004). How fading work solution steps works – A cognitive load perspective. *Instructional Science, 32*, 59-82.

Robins, S. & Mayer, R.E. (1993). Schema Training in Analogical Reasoning. *Journal of Educational Psychology, 85(3),* 529-538.

Roy, M., & Chi, M. T. H. (2005). The self-explanation principle in multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 271–287). Cambridge, England: Cambridge University Press.

Salomon, G. & Perkins, D.N. (1989). Rocky roads to transfer: rethinking mechanisms of a neglected phenomenon. *Educational Psychologist, 24,* 113 – 142.

Streveler, R.A., Nelson, M.A., Ronald L. Miller, R.L., Olds, B.M., Evans, D.L., Mitchell, J., & Martin, J. (2004). Investigating the conceptual understanding of engineering students. *Proceedings of Annual Conference of the American Educational Research Association*, San Diego, CA, April, 2004.

Streveler, R. A., Litzinger, T.A., Miller, R.L., & Steif, P.S. (2008). Learning conceptual knowledge in the engineering sciences: Overview and future research directions. *Journal of Engineering Education*, *97(3)*, 279-294.

Sweller, J. (2005). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 159–168). Cambridge, England: Cambridge University Press.

Sweller, J. & Cooper, G.A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction, 2*(1), 59-89.

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

Schworm, S., & Renkl, A. (2006). Computer-supported example-based learning: When instructional explanations reduce self-explanations. *Computers and Education, 46,* 426–445.

Schworm, S., & Renkl, A. (2007). Learning argumentation skills through the use of prompts for self-explaining examples. *Journal of Educational Psychology, 99,* 285–296.

Tarmizi, R.A. & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology, 80*(4), 424-436.

Uttl, B. (2005). Measurements of individual differences: Lessons from memory assessment in research and clinical practice. *Psychological Science, 16*(6), 460–467.

van Gog, T., Ericsson, K.A., Rikers, R.M.J.P., & Paas, F. (2005). Instructional design for advanced learners: Establishing connections between the theoretical frameworks of cognitive load and deliberate practice. *Educational Technology, Research, and Development, 53(3),* 73-81.

van Gog, T., Paas, F., & Van Merrienboer, J. J. G. (2006). Effects of process-oriented worked examples on troubleshooting transfer performance. *Learning and Instruction, 16(2),* 154-164.

Vygotsky, L.S. (1978). *Mind and society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Ward, M. & Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction, 7*(1), 1-39.

Wittwer, J. & Renkl, A. (2008). Why instructional explanations often do not work: A framework for understanding the effectiveness of instructional explanations. *Educational Psychologist, 43(1),* 49-64.

Wood, D., Bruner, J.S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 17,* 89–100.

Zimmerman, B. (1999). Attaining self-regulation: A social cognitive perspective. In P. R. P. Boekaerts, & M. Zeidner (Ed.), *Handbook of self-regulation* (pp. 13-41). San Diego, CA: Academic Press.

APPENDIX A Informed Consent Form

This IRB is used for the complete CE 360 project, and involves some of activities that are not a part of this study.

INFORMED CONSENT FORM FOR SOCIAL SCIENCE RESEARCH
The Pennsylvania State University

Title of Project: Investigation of the Use of Tablet PC for Problem-Solving in Engineering Courses

Principal Investigator: Roxanne Toto, rtoto@psu.edu
201 Hammond Building
University Park, PA 16802
(814) 865-4016

1. **Purpose of the Study**: The purpose of this study is to examine your perceptions of the problem-solving activity in this course. In particular, we want to understand what you like and don’t like about the problem-solving activity. The information collected in this project will be used to improve future sections of engineering courses.

2. **Procedures to be followed:**
   - If you agree to take part in this research, you will be asked to complete a survey in class at the end of the semester. It should take no more than 20 minutes. The survey will ask questions about your perceptions of the problem-solving activities in the course. Some of you will be asked to participate in the focus group or possibly an interview at the end of the semester. The focus group will take less than 40 minutes. The interview will take no more than 30 minutes. Food and refreshments will be provided during the focus group.
   - If you agree to take part in this research, you are also giving the researchers permission to use your grades on homework assignments and quizzes.

3. **Benefits:**
   - You will benefit because the data will be used to make improvements to your class. The information collected will also help to improve future development of problem-solving activity.
   - This research will add to the pool of knowledge about technology use in engineering classrooms.

4. **Duration/Time:** It will take no more than 20 minutes to complete the survey. If you participate in the focus group, it will take no more than 40 minutes. If you participate in the interview, it will take no more than 30 minutes.

5. **Statement of Confidentiality:** Your participation in this research is confidential. Only the person in charge, and his/her assistants, will know your identity. The data will be stored and secured at 201 Hammond Building in a locked cabinet and/or password-protected electronic file. In the event of a publication or presentation resulting from the research, no personally identifiable information will be shared. Only the investigators listed above will have access to personally identifiable information. All identifying information will be removed prior to data analysis. For data collected online, your confidentiality will be kept to the degree permitted by the technology. No guarantees can be made regarding the interception of data sent via the Internet by any third parties.

6. **Right to Ask Questions:** You may ask questions about the research at any time by contacting Roxanne Toto, e-Learning Support Specialist at the Leonhard Center for the Enhancement of Engineering Education, 201 Hammond Building, 814-865-4017, rtoto@psu.edu

7. **Voluntary Participation:** Participation is voluntary. You can withdraw from the study at any time by notifying the principal investigator. You can decline to answer specific questions.

8. **Audio Recording:** Audio recording may take place during the focus group and the interview to ensure accuracy. Audio records will be stored in 201 Hammond and destroyed by June 30, 2011. Access to the audio recording will be limited to the individuals names as investigators and research personnel (Roxanne Toto, Sarah Zappe, Thomas Litzinger, Hien Nguyen, and Natalia Kapli)

You must be 18 years of age or older to consent to participate in this research study.

If you agree to the conditions and statements noted above, please provide your signature and the date below.

☐ I agree to allow my grades (final course grade, homework assignments, and class activities) for CE 360 to be released to the principal investigator and the research team of this study for the purpose of understanding the impact of active learning activities in the classroom.

☐I DO NOT to allow my grades (final course grade, homework assignments, and class activities) for CE 360 to be released to the principal investigator and the research team of this study for the purpose of understanding the impact of active learning activities in the classroom.

Please complete the section below:

Student Name Signature

Date

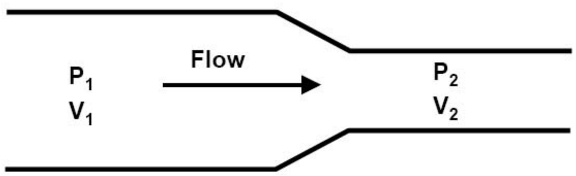
Investigator Name Signature

Date

APPENDIX B Conceptual Knowledge Instrument

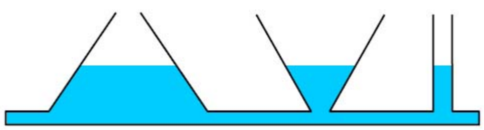
Conceptual Knowledge Instrument (displayed with permission)

1. Water flows through a pipe and enters a section where the cross sectional area is smaller. Viscosity, friction, and gravitational effects are negligible. Select the correct statement about the change in pressure P and average velocity V.

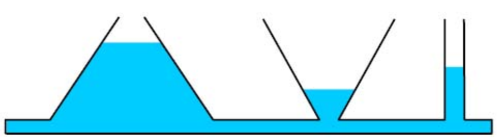


A) P2 is less than P1 and V2 is less than V1
B) P2 is less than P1 and V2 is greater than V1
C) P2 is greater than P1 and V2 is less than V1
D) P2 is greater than P1 and V2 is greater than V1

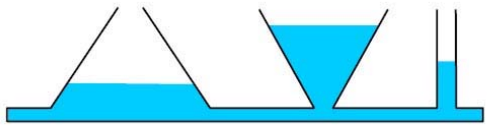
2. Three containers connected at the base are filled with a fluid. The top of each container is open to the atmosphere and surface tension is negligible. The container shapes are all different. Select the figure that shows the correct fluid levels in the containers at equilibrium conditions.

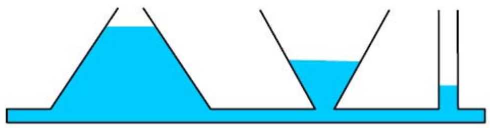


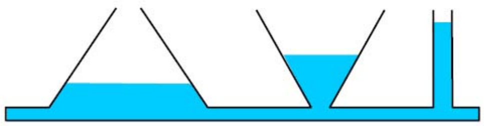
A)



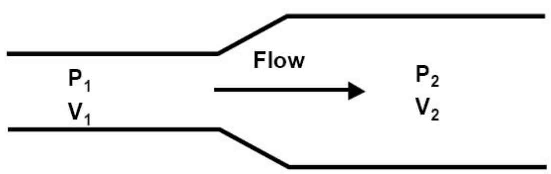
B)

C) 

D) 

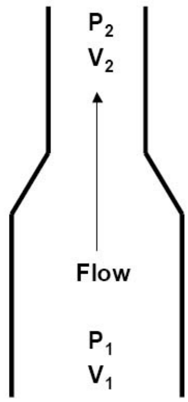
E) 

3. Water flows through a pipe and enters a section where the cross sectional area is larger. Viscosity, friction, and gravitational effects are negligible. Select the correct statement about the change in pressure P and average velocity V.



A) P2 is less than P1 and V2 is less than V1
B) P2 is less than P1 and V2 is greater than V1
C) P2 is greater than P1 and V2 is less than V1
D) P2 is greater than P1 and V2 is greater than V1

4. Water flows vertically up through a pipe and enters a section where the cross sectional area is smaller. Viscosity and pipe friction effects are negligible but gravitational effects are not negligible. Select the correct statement about the pressure P2 and velocity V2.



A) P2 equals P1 and V2 equals V1

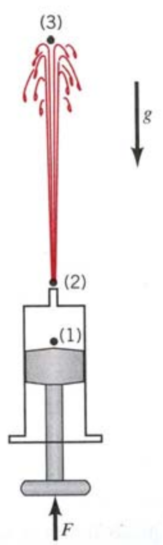
B) P2 is greater than P1 and the V2 is greater than V1

C) P2 is greater than P1 and the V2 is less than V1

D) P2 is less than P1 and the V2 is greater than V1

E) P2 is less than P1 and the V2 is less than V1

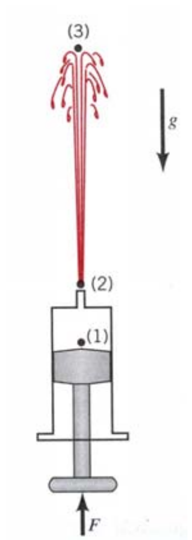
5. Consider the flow of water from the syringe. At which location(s) is kinetic energy the greatest?



A) Point (1)
B) Point (2)

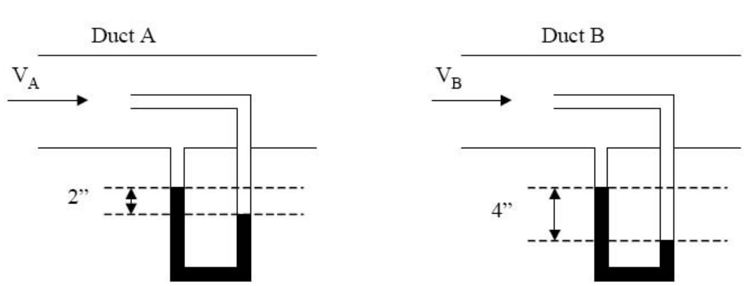
C) Point (3)
D) Points (2) and (3)

6. Consider the flow of water from the syringe. At which location(s) is pressure the greatest?



A) Point (1)
B) Point (2)
C) Point (3)
D) Points (2) and (3)

7. Pitot tubes are placed in two ducts in which air flows as shown below. The density and temperature of the flows are equal. The dynamic (velocity) pressure and the static pressure taps are connected to two manometers. The pressure difference for duct A is 2" of water and that for Duct B is 4" of water. Select the correct answer for the velocity VA in duct A relative to the velocity VB in duct B.



A) VB equals 2 VA

B) VB equals $\sqrt{2}$VA

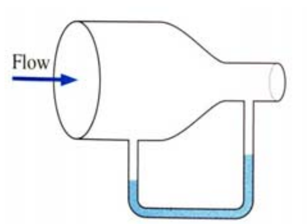
C) VB equals VA

D) VB equals VA / $\sqrt{2}$

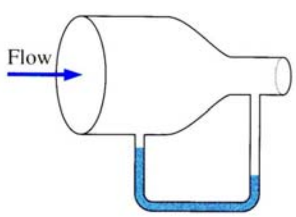
E) VB equals VA / 2

8. A glass manometer with oil as the working fluid is connected to an air duct. Will the oil in the manometer move as in option A or B, or remain on the same level as in option C?

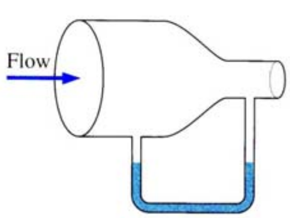
A)



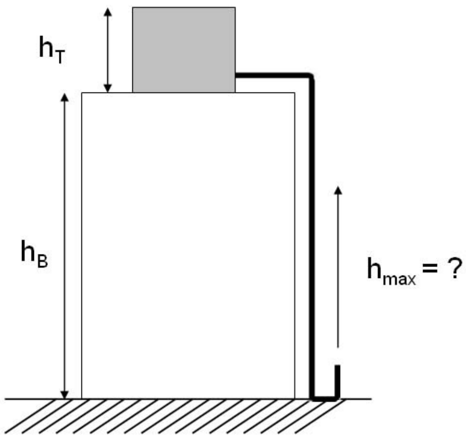
B)



C)



9. The water level of a tank on a building roof is a total of hB + hT meters above the ground. A hose leads from the tank bottom to the ground. The end of the hose has a nozzle, which is pointed straight up. Friction effects are negligible. Select the correct statement about the maximum height to which the water could rise.



A) $h_B / \sqrt{2}$

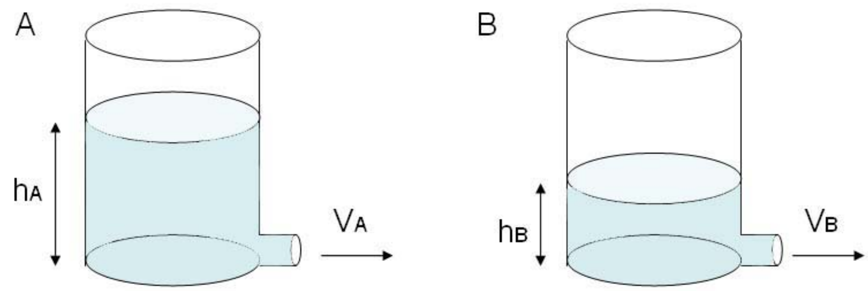
B) h_B

C) $h_B + h_T$

D) $(h_B + h_T) / \sqrt{2}$

E) $(h_B + h_T) / 2$

10. Two large tanks open to atmosphere are filled with water to heights hA and hB from the outlet tap. hB is 1 / 2 of hA. A tap near the bottom of each tank is now opened, and water flows out from the smooth and rounded outlet. Select the correct statement about the relationship between VA and VB.



A) VB equals 2 VA

B) VB equals $\sqrt{2}$VA

C) VB equals VA

D) VB equals VA / $\sqrt{2}$

E) VB equals VA / 2

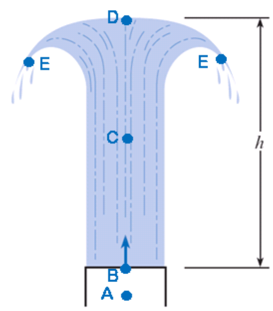
APPENDIX C Problem Solving Instrument

Problem Solving Instrument (displayed with permission)

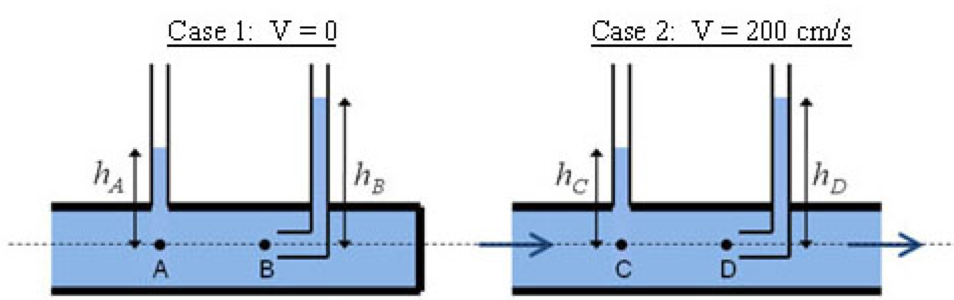
1. If you were given the flowrate or velocity out of the pipe and asked to compute h, at which two locations would you choose to apply the Bernoulli equation and why?

Location 1: _______, because

Location 2: _______, because



2. Base your answer to question 2 on the diagram below, which shows two pipe systems with different tubes for measuring energy in the water. In case 1, the water is static and the velocity equal to zero, whereas in case 2, the water is flowing through the pipe at a velocity of 200 cm/s. Both pipes have the same diameter, the pressures at points A and C are equal, and points A, B, C and D are all located at the same datum. Note that the h_A, h_B, h_C and h_D are **not drawn to scale**.



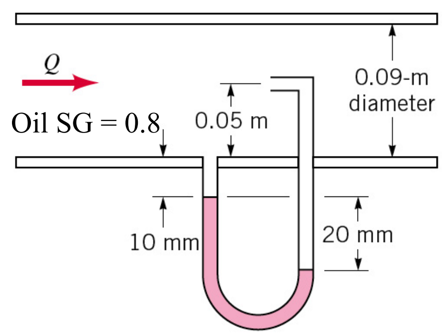
Given that $p_A/\gamma = p_C/\gamma = 0.05$ m, you can calculate h_B and h_D, which have values of _____ and _____.

a) 0.15 m; 0.15 m
b) 0.05 m; 0.10 m
c) 0 m; 0.10 m

d) 0.05 m; 0.25 m

3. Oil of specific gravity 0.80 flows in the pressurized constant-diameter pipe shown below. Neglecting head losses, calculate the flow through the pipe. The manometer fluid is water.

Note: Unlike the similar homework problem you did online, here the fluid is oil (not air), so you must take into account the static fluid pressure as it varies with depth.



Natalia V. Kapli
VITA

Education

2010	Ph. D. in Instructional Systems, The Pennsylvania State University Minor in Workforce Education
2001	M.S. in Adult Learning, Performance, & Development, Drake University, IA
1996	Diploma (B.A.) in English and German Teaching, Novosibirsk State Pedagogical University, Novosibirsk, Russia

Employment

2007 - 2010	Research Assistant / Instructor Assessment & Evaluation, Education Technology Services, Penn State University Engineering Assessment & Instructional Support, Penn State University
2006	Learning & Education Intern PricewaterhouseCoopers
2005	Instructional Designer/Technical Designer – Intern Bank of America
2003 - 2005	Graduate Assistant Engineering Assessment & Instructional Support, Penn State University Multimedia Lab, Instructional Systems, Penn State University
2002 - 2003	Research Associate – CCNMTL Fellow Columbia Center for New Media Teaching & Learning (CCNMTL), Columbia University

Selected Publications

Litzinger, T., Van Meter, P., **Kapli, N.**, Zappe, S., & Toto, R. (2010). Translating education research into practice within an engineering education center: Two examples related to problem solving. *International Journal of Engineering Education, 26(4),* 860-868.

Nguyen, H., **Kapli, N.,** & Ching, Y. (2008, November). Facilitating the community of inquiry in a large online introductory astronomy course. *E-Learn Proceedings,* Las Vegas, NV.

Kapli, N., Wise, J.C., Litzinger, T.A., & Donohue, W. (2006, June) Preparing students to compete in the global marketplace. *Proceedings of the American Society for Engineering Education (ASEE) Annual Conference and Exhibition*, Chicago, IL.