The Design and Utilization of Effective Worked Examples:

A Meta-Analysis

By

Jennifer K. Crissman

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Educational Studies

Specialization: Educational Leadership and Higher Education

Under the Supervision of Professor Jody C. Isernhagen

Lincoln, Nebraska

May, 2006

UMI Number: 3208114

The Design and Utilization of Effective Worked Examples:

A Meta-Analysis

Jennifer Karina Crissman, Ph.D.

University of Nebraska, 2006

Advisor: Jody C. Isernhagen

The use of worked examples represents an alternative instructional method to that of problem-solving in highly structured domains such as physics and computer programming. The benefits of worked examples may be measured in posttest performance scores, and lower cognitive load in the form of less perceived mental effort. However, for worked examples to be effective, research into the range of different design components is crucial for creation of effective worked examples. Such components include intra-example features, such as the integration of source information with diagrammatic or auditory material, labeling or cueing of problem subgoals, and faded solution steps, and training in self-explanation strategies. Inter-example features such as the presentation of multiple examples, using example/problem pairs, and varying surface stories and problem types, are also important components of effective worked examples. A meta-analysis of multiple studies that employ various forms of worked examples instructional techniques was conducted. Mental effort ratings and posttest performance scores were used to calculate the standardized mean difference effect sizes for each study.

The findings of this study showed that worked examples indeed represent a benefit to learning, with certain features and/or environments producing greater benefits than

others. When the general analysis was broken down into intra-example features and self-explanation training, the results showed that faded solution steps demonstrated the strongest effect in favor or worked example instruction, followed by self-explanation training. Conventional worked examples, integrations, and subgoals also represented moderate effects. Inter-example features such as presenting multiple examples per problem type and surface story variation produced better outcomes in terms of posttest scores compared to when these features were not present. Prior knowledge was also a factor in the efficacy of worked examples as students with little to no prior knowledge benefited more than those that had prior knowledge in the domain.

Acknowledgements

The completion of this dissertation is culmination of a challenging and rewarding doctoral program. It is upon completion of my research that I find this page the most challenging to write. There are not enough words to thank all those who gave their support and encouragement.

First I would like to thank my advisor, Dr. Jody Isernhagen, whose tireless help and guidance made the completion of this dissertation possible. I am extremely thankful for her willingness to let me pursue this type of research and her faith that I would get it done right. It was a tremendous pleasure to be advised by her.

I would like thank the members of my committee, Dr. Dlugosh, Dr. Grady, Dr. Joekel, and Dr. Ansorge. I truly appreciate their invaluable recommendations and assistance throughout this process.

To my friends and colleagues who read my paper when it was the last thing they wanted to do and who always reminded me that there was a light at the end of the tunnel (and it wasn't a train).

To my family who never failed to encourage me during the low times and always pretended to be interested in my work the rest of the time. I am thankful to my parents for their support and obvious pride in having the "smartest member in the family." Thanks to my mother and father in-law for their constant encouragement. And, thanks to Grandpa who always believed that I would finish and called me "Doctor" from the beginning.

Finally, my heartfelt thanks go to my husband and son for their patience, love, and support. To Greg for his understanding, for talking me through the frustrations, and for taking care of everything, including me, without complaint. And to Tyler for making me laugh every time I needed it.

TABLE OF CONTENTS

Page

Page

Page

Page

Page

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

APPENDIX PAGE

Chapter I

Introduction

Background

Since the 1950s and 1960s, the dominant method of instruction for many highly structured fields, such as mathematics, physical sciences, and computer science, has been discovery learning in the form of problem-solving (Tuovinen, 1997). Problem-solving is described as a "process coordinating previous experiences, knowledge, and intuition to determine an outcome in a situation where such a procedure is not known" (Lester & Knoll, 1990). It is often used synonymously with discovery learning (Zhu & Simon, 1987). This instructional technique involves very little structure and guidance, but is often the method of instruction employed in highly structured disciplines (Tuovinen & Sweller, 1999). Problems arise in discovery learning when students misunderstand the problem and/or the information given to them. Students must be able to assemble all of the given information and what is being asked of them into their working memory and then match these with the knowledge structure and the repertoire of problem schemas stored in the long term memory (Lai, Griffin, Mak, Wu, & Dulhunty, 2001). Criticisms of discovery learning include: its lack of generalizability, the possibility of imposing a high cognitive load, and the substantial need of the student for continuing guidance from the instructor (Tuovinen & Sweller, 1999).

Worked examples provide an alternative to problem-solving. Worked examples consist of a problem statement and the appropriate steps to the solution (Kalyuga,

Chandler, Tuovinen, & Sweller, 2001). The use of worked examples in instruction traces its origins to cognitive load theory.

Sweller's (1988) cognitive load theory is one explanation for the efficiency of worked examples over problem-solving. This theory assumes a very limited working memory, and an unlimited long-term memory which holds many schemas (Tuovinen & Sweller, 1999). Schemas represent procedures for problem-solving activities and vary in the level of how they are automatically applied (Tuovinen & Sweller, 1999). Recent studies have shown that worked examples may have an advantage over problem-solving, at least for initial skill acquisition and reduction of cognitive load (Carroll, 1994; Lee, Nicoll, & Brooks, 2004; Renkl & Atkinson, 2003; Tuovinen & Sweller, 1999).

Care should be taken when utilizing worked examples as there are instances where they can represent a detriment to learning. Research has shown that certain factors contribute to the effectiveness of worked examples. One factor includes intra-example features or how the examples are designed with emphasis on how the solution is presented (Atkinson, Derry, Renkl, & Wortham, 2000). Additionally, inter-example features, or how multiple examples relate to each other and practice problems within a lesson, play a role in the efficacy of worked examples (Atkinson et al., 2000). Finally, individual differences in how examples are processed, especially in terms of how learners explain solutions to themselves, also play a critical role in the worked example effectiveness (Atkinson et al., 2000).

Problem Statement

This study concentrated on the effectiveness of worked examples for students of higher education, specifically examining subjects ranging from high school students to

adult learners, as representative of a community college population. The research on worked examples has examined their efficacy on subjects ranging from elementary students (Chi, De Leeuw, Chiu, & LaVancher, 1994; Mousavi, Low, & Sweller, 1995; Zhu & Simon, 1987) to the elderly (van Gerven, Paas, van Merrienboer, & Schmidt, 2002; van Gerven, Paas, van Merrienboer, Hendriks & Schmidt, 2003).

Almost half, 44%, of American undergraduates attend community colleges (American Association of Community Colleges, 2000). Students attending community colleges do not necessarily typify the “traditional” 18-22 year-old, full-time college students. While the largest group attending community colleges are typical students, there is a large population of adult learners, that is those that are older than 25 (Howell, 2003). According to the American Association of Community Colleges (Patton, 2000), almost 60% of those students attending community colleges are above the age of 22, with the average age being 36. In addition to adult learners, community colleges often enroll high school students as part of secondary education programs, such as the Post-Secondary Enrollment Option (PSEO) program (Taft, 2005). This diverse mixture of students makes studying the efficacy of worked examples extremely important.

The research pertaining to subjects ranging from high school students to adult learners holds contradictions. An investigation of relevant research corresponding to the use of worked examples would provide insight into instructional design issues and benefits.

Statement of Purpose

The purpose of this meta-analysis was to examine worked examples as compared to other instructional methods by determining their general efficacy in addition to analyzing

the components of efficient worked examples in terms of their structure, inducement of student interaction, and relationship to previous knowledge. Problem-solving has been the traditional pedagogical method for teaching in such domains as mathematics and physical sciences. Studies on cognitive load suggest that worked examples can produce more effective learning outcomes, especially for near transfer. While problem-solving skills are the ultimate goal in structured disciplines, worked examples provide an intermediary tool to gain these skills. Much research has been conducted on the efficacy of worked examples but its synthesis had not yet been conducted. By conducting a meta-analysis of the literature, resolutions can be made regarding contradictory findings, relationships between learning outcomes and worked examples exposed, and avenues of future research identified.

Research Questions

This study quantitatively analyzed the general efficacy of worked examples and components of effective worked example instruction. The research investigated the effective design and utilization of ideal worked examples as compared to conventional instruction in terms of cognitive load. This comparison was based on intra-example features that deal with how examples are designed (Catrambone, 1994; Catrambone & Holyoak, 1990; Mwangi & Sweller, 1998; Ward & Sweller, 1990), and inter-example features which deal with how examples are utilized and presented (Catrambone & Holyoak, 1989; Paas & van Merrienboer, 1994; Quilici & Mayer, 1996; Reed & Bolstad, 1991). In addition this study analyzed the effects of self-explanation training with examples (Chi, et al., 1994; Renkl, Stark, Gruber, & Mandl, 1998), and the effects of previous knowledge of the learner on the utility of worked examples (Chi, Feltovich, &

Glaser, 1981; Kalyuga, Chandler, & Sweller, 2001; Kalyuga, Chandler, Tuovinen, et al., 2001). These design and presentation features as well as previous knowledge and self-explanation training were further examined in relation to their effect on the cognitive load of the leaner (Cooper & Sweller, 1987; Sweller & Cooper, 1985).Consequently, this study was directed by these general categories: (a) intra-example features, (b) inter-example features, (c) self-explanations, (d) previous knowledge, and (e) cognitive load.

This study is guided by the following research questions:

1. Do worked examples have an overall significant advantage over conventional instructional techniques?
2. Is the efficacy of worked examples as compared to conventional instructional methods affected by the integration of source materials, either visually, aurally, or a combination?
3. Is the fading of solution step within worked examples a beneficial technique in terms of learning outcomes as compared to conventional instruction?
4. Do inter-example features such as multiple examples, variability of problem types, variability of surface stories and example-problem pairs make a difference in learning outcomes of high school to adult learners?
5. Does training subjects in self-explanation techniques result in positive learning outcomes for subjects ranging from high school students to adult learners?
6. Does the previous knowledge of subjects ranging from high school students to adult learner students have an inverse relationship on the efficacy of worked examples with their effectiveness being reduced by subjects with more prior knowledge?

7. Do worked examples serve to reduce cognitive load for subjects ranging from high school students to adult learners?

Theoretical Base

Sweller's (1988) cognitive load theory permeates much of the research on worked examples. Cognitive load theory presents the architecture of the human mind as having limited working memory and virtually unlimited long-term memory that holds large numbers of schemas that can vary in their degree of automation (Tuovinen, 1997). This theory relates to instructional materials used due to the individual differences of the learner, the characteristics of the actual material, and the interaction between the learner and material. A learner with the schemas in place to learn a new concept will have significantly less cognitive load than a learner who does not, and subsequently much less working memory will be used to translate the material to long-term memory by incorporating it into a schema. Characteristics of the instructional material incorporate the interrelatedness of the elements. The higher the organization of elements, the higher the cognitive load. Finally, material that imposes a heavy cognitive load for some people because they must deal with large numbers of interacting elements may impose less cognitive load for other people because they have acquired automated schemas that incorporate the individual elements (Tuovinen & Sweller, 1999).

Effective instructional materials should attempt to minimize activities that do not directly encourage learning in order to prevent the overload of working memory capacity, as indicated by cognitive load theory (Renkl, Atkinson, & Grobe, 2004). Three types of load that are described by Sweller, van Merrienboer, and Paas (1998) should be taken into account when designing instructional materials. They are intrinsic load, which is due

to the nature of the material, extrinsic load which is due to unrelated activities to the learning process, and germane load, or mental activities that are relevant to learning. These three types of load are interconnected, in that care should be taken not to induce high extraneous load, especially when coupled with intrinsic load, while attempting to promote germane load (Renkl et al., 2004; Sweller et al., 1998)

The assertion made by cognitive load theory is that optimal learning conditions exist when learning is aligned with human cognitive architecture. This encourages: 1) incorporating worked examples into instructional materials, 2) reducing, or even eliminating the cognitive load on working memory that is associated with having to mentally integrate several sources of information, and 3) increasing the capacity of working memory by such practices as chunking or using auditory as well as visual information under conditions where both sources of information are essential to understanding (Yang, 2002).

Although cognitive load theory has been used to explain why worked examples are more effective than certain forms of problem-solving, it can also be used to explain why worked examples can be ineffective. Such an instance occurs when the components of an example, such as text and diagrams, result in a split-attention effect and must be mentally integrated by the learner, thus increasing cognitive load (Tarmizi & Sweller, 1988).

Sweller, Chandler, Tierney and Cooper (1990) developed a six-point theory to aid in the construction of instructional materials based on prevention of the split attention effect, and cognitive load theory.

Cognitive laod theory is described as follows:

1. Schema acquisition is the building block of skilled performance,
2. Schema acquisition requires attention directed to problem states and their associated moves; other cognitive activities must remain limited and peripheral so as not to impose a heavy cognitive load that interferes with learning,
3. Encouraging or requiring learners to engage in means-end search or to integrate multiple sources of information misdirects attention and imposes a heavy cognitive load,
4. Because integrating multiple sources of information misdirects attention and imposes a heavy cognitive load, schema acquisition cannot commence until disparate sources of information have been mentally integrated,
5. Material with reduced or unitary sources of information will reduce or eliminate the need for learners to use cognitive resources to restructure material into a form suitable for schema acquisition, and
6. Learning is enhanced when learners are allowed to attend to schema acquisition rather than to information reformulation (Sweller et al., 1990).

In order to alleviate the split-attention effect, and resulting cognitive load, of some instructional materials, Mousavi et al. (1995) investigated the ramifications of the modality effect. This theory suggests that presenting materials in a dual-mode format, such as aurally and visually, may help to reduce cognitive load and increase processing capability (Jeung, Chandler, & Sweller, 1997; Mousavi et al, 1995.). Mousavi et al. (1995) went on to propose that current theories on memory assume the existence of multiple working memory stores rather than a unitary one. These memory stores are

commonly used for auditory and visual processing (Mousavi et al., 1995). They stated that if these two stores work independently, the size of working memory can be increased, and consequently increasing the amount of information that can be processed by using a dual-mode presentation (Mayer (1997, 2001; Mousavi et al., 1995) extended the modality theory to multimedia learning.

According to cognitive load theory (Sweller, 1988; Sweller et al., 1998) and Mayer's (2001) theory of multimedia learning, replacing visual text with spoken text (*the modality effect*) and adding visual cues relating elements of a picture to the text (*the cueing effect*) both increase the effectiveness of multimedia instructions in terms of better learning results or less mental effort spent.

Mayer (2001) further asserted that due to the limited capacity of working memory, cognitive load could be reduced for multimedia presentations by encouraging learner interactivity. This allows learners to digest and integrate one segment of the explanation before moving on to the next, creating an interactivity effect.

In addition to structuring worked examples to include a dual-mode presentation format, the organization of knowledge around subgoals, or meaningful parts of an overall solution procedure (Catrambone, 1998), have proved beneficial.

Atkinson, Catrambone, and Merrill (2003) developed the subgoal learning model which can be summarized as follows:

1. By using a cue, such as segregating a set of grouped steps or labeling them, learners are encouraged to treat a collection of solution steps as a group,
2. After grouping the steps, learners typically attempt to explain to themselves why those steps go together in order to determine its purpose, and

3. Through the process of self-explanation, learners form a goal that captures the purpose of that set of solution steps (Atkinson, Renkl, & Merrill, 2003).

The self-explanation effect was described by Chi, Bassok, Lewis, Reimann, and Glaser (1989) in terms of the effectiveness of worked examples being related to how students explain solutions to themselves. They asserted that good students will succeed in generalizing because he or she makes a conscious effort to ascertain the conditions of application of the solution steps beyond what is explicitly stated. To do so, the student must "explain" how the example instantiates the principle which it exemplifies (Chi et al, 1989). They suggest that explanations can reveal students' understanding by showing whether or not they know: 1) the conditions of application of the actions, 2) the consequences of actions, 3) the relationship of actions to goals, and 4) the relationship of goals and actions to natural laws and other principles (Chi et al., 1989).

Paas and van Merrienboer (1993, 1994) developed a procedure for calculating instructional effectiveness in terms of two factors: perceived mental effort and performance. This procedure involves transferring the mental effort and performance data to z-scores and displaying it as an M (mental effort) – P (performance) grid (Paas & van Merrienboer, 1994). Efficiency can be determined by the relation between mental effort and performance in a particular condition in relation to a hypothetical baseline condition (Paas & van Merrienboer, 1994). The high efficiency portion of the grid would consist of low mental effort and high performance scores, while conversely the low efficiency portion would be characterized by high mental effort and low performance scores. This procedure for calculating instructional effectiveness has been used in many subsequent