8-8-2009

The effect of a worked example on online debate quality in an information systems course

Christopher Scott Tollison

THE EFFECT OF A WORKED EXAMPLE ON ONLINE DEBATE QUALITY IN AN INFORMATION SYSTEMS COURSE

By

Christopher Scott Tollison

A Dissertation
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in
Instructional Systems and Workforce Development
in the Department of Instructional Systems
and Workforce Development

Mississippi State, Mississippi

August 2009

THE EFFECT OF A WORKED EXAMPLE ON ONLINE DEBATE QUALITY

IN AN INFORMATION SYSTEMS COURSE

By

Christopher Scott Tollison

Approved:

Kui Xie
Assistant Professor of Instructional Systems and Workforce Development
(Major Professor and Director of Dissertation)

James Adams
Associate Professor of Instructional Systems and Workforce Development
(Committee Member)

Connie M. Forde
Professor of Instructional Systems and Workforce Development
(Committee Member)

Jianzhong Xu
Professor of Leadership and Foundations
(Committee Member)

Linda F. Cornelious
Professor and Interim Department Head of Instructional
Systems and Workforce Development
(Graduate Coordinator)

Richard Blackbourn
Professor and Dean of the College of Education

Name: Christopher Scott Tollison

Date of Degree: August 8, 2009

Institution: Mississippi State University

Major Field: Instructional Systems and Workforce Development

Major Professor: Dr. Kui Xie

Title of Study: THE EFFECT OF A WORKED EXAMPLE ON ONLINE DEBATE QUALITY IN AN INFORMATION SYSTEMS COURSE

Pages in Study: 136

Candidate for Degree of Doctor of Philosophy

The current study investigates the effects of preparing learners for an online debate through a worked example in terms of student perception, participation, level of cognitive skill, and electronic interaction patterns.

There has been a change in the focus of distance learning research from comparative media studies to the means to improve the quality of distance education. One of the key elements in this changing impetus are strategies to promote interaction such as the introduction of structure or scaffolding argumentation (McIsaac & Blocher, 1998). One such strategy is the online debate in which students are organized into teams to take a position on an issue and argue on its behalf (Jeong, 2004). The debate is constrained through the addition of rules and specific message headers.

Well-designed quality interaction holds the potential to create more satisfied learners and higher quality learning outcomes (Muirhead, 2002), but in the absence of

quality, interaction has been found to actually lead to a decrease in satisfaction, participation, and performance (Joung & Keller, 2004; Kreijns, Kirschner, & Jochems, 2002). One way to ensure quality within interactive exercises such as online debate is to prepare students through a worked example. A worked example models an expert's work and demonstrates desired behaviors for the learner to study (Atkinson, Derry, Renkl, & Wortham, 2000).

Students were randomly assigned to teams to participate in an online debate with half being given access to a worked example before participating. In order to examine the effects of the worked example on students' perceived satisfaction and level of preparedness, a survey was administered at various points throughout the semester. Additionally, debate transcripts were analyzed for participation, cognitive skill, and interaction patterns. The results demonstrate that students prepared through a worked example participated more frequently, wrote more words or phrases that encouraged the participation of others, and used higher-order thinking skills. The conclusion was that worked examples can be used to model behaviors for students to emulate. The implication being that instructors should consider providing worked examples before engaging students in online debate and future research should examine the efficacy of a worked example in preparing learners for other types of interactive activities.

DEDICATION

This dissertation is dedicated to my mother and father whose sacrifices on behalf of this endeavor are too numerous to list. Even as I type this, my mom is probably hard at work performing some task that I should be doing for myself and my dad is likely framing a doorway or hanging a piece of drywall at my home. It seems unfair that only one name goes on the byline because this dissertation is as much the result of their labor and their devotion as it is my own.

ACKNOWLEDGEMENTS

It is with a tremendous amount of gratitude that I wish to acknowledge those who made this project possible.

I would like to thank my friends and particularly my family who provided the personal encouragement and motivation necessary to sustain me throughout this process.

I would also like to acknowledge my colleagues within the College of Business and Legal Studies at the Mississippi University for Women who have provided me both the flexibility and support to put in the long hours necessary to complete an undertaking of this scale.

I would also like to offer my sincere appreciation to my first dissertation committee chair Dr. Vance Durrington whose assistance in developing the initial vision for this project helped me round first base, my current dissertation committee chair Dr. Kui Xie whose guidance, enthusiasm, and dedication has proven invaluable in moving this dissertation across home plate, and committee members, Dr. James Adams, Dr. Connie Forde, and Dr. Jianzhong Xu, for their tireless efforts not only on behalf of this project, but also as top-notch educators and mentors for so much of my doctoral coursework. I am far better for having received their counsel.

I would also like to recognize the faculty at Mississippi State University. In completing three degrees and spending over a decade on campus, I have had the opportunity to take courses under the tutelage of an untold number of faculty members.

It is without a moment's hesitation that I can say I have learned something from and been personally enriched by each and every one of them.

I would also like to recognize the staff at the University (many of whom I have bothered on far too many occasions). Their hard work in support of the University mission too often goes unnoticed, but nonetheless is a critical element of the foundation on which any great institution is built. They are among the many reasons I shall forever be proud to wear Maroon and White.

TABLE OF CONTENTS

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURE

CHAPTER I

INTRODUCTION

Distance learning has seen tremendous growth in recent years. A 2009 report of the National Center for Education Statistics indicated that 66% of 2-year and 4-year Title IV eligible, degree granting institutions offered distance education courses during the 2006-2007 academic year and that over 12 million students were enrolled in a distance education course. The asynchronous Internet-based course spawned by the popularity of the World Wide Web has been most responsible for the recent growth in distance learning (Vamosi, Pierce, & Slotkin, 2004). In 2006-2007, three-quarters of all institutions utilizing distance education reported using the asynchronous online course as a primary technology for delivering distance education courses – a figure six times greater than the next most utilized technology (NCES). Furthermore, 98% of institutions offering distance education courses reported using asynchronous online technologies to some extent.

Muilenburg and Berge (2000) reported that the majority of distance education courses rely on some type of online forum. Although the online forum is not a real-time activity, it still takes these courses beyond being relatively simple exercises in note printing, task posting, and assignment submitting. First, these online discussion systems encourage the participation of students scattered over great distances and at different times. Beyond breaking down the face-to-face boundaries of time and distance, Hannafin (1999) found that online discussion also enhances student learning due to its emphasis on reflection and higher-order thinking. These findings about online forum and online discussions are confirmed by broader findings in the field of computer-supported collaborative learning (CSCL). Kreijns, Kirschner, and Jochems (2002) recognized the practical benefits of CSCL like anywhere-anytime learning and cost efficiencies due to the need for fewer instructors, but also cited the instructional benefit of generating group learning. Joung and Keller (2004) associated CSCL with five benefits – academic achievement, increased levels of student satisfaction, better individual and group products, better group cognition, and like Hannafin before, greater higher-order thinking skills.

Statement of the Problem

Research about distance education has been ongoing for decades. Historically, these studies have attempted to gauge the effectiveness of using different delivery formats and technologies and have shown no demonstrable difference in student performance. With more faculty members teaching at a distance and more students enrolling in these courses, there seems to have been an emergence in the number of instructional methods. An example would be having students debate one another on online discussion boards (i.e. online debates). These factors have collectively provided a greater impetus for a new direction in distance learning research (Lee, Driscoll, & Nelson, 2004; McIsaac & Blocher, 1998).

This new direction is not geared towards the "no significant difference phenomenon" and proving the relative merit of distance learning, but instead it is geared to ensuring the best learning environment possible for the distance learner including issues such as transactional distance, learner control, the social dimension, and strategies to increase interactivity (McIsaac & Blocher, 1998). Examination of these four interrelated areas could shed light on the reason that some researchers identified end results of low participation, low quality learning, and lower levels of satisfaction (Kreijns et al., 2002) with distance education and other computer supported learning environments. A possible reason for a less productive learning environment is the lack of quality interaction brought on by – as one researcher stated – instructors simply assuming that social interaction will work (Kreijns et al.). Consequently, a further exploration of the quality of the interaction may be considered a critical element in online course development. In many cases, the activities designed to promote interactivity within courses prove difficult to students.

Although online debates are designed with the promise of more interaction, their stringent adherence to rules and their hypertext context makes them more challenging for the students which, in turn, may decrease the amount or quality of their interaction. As a possible solution, Jonassen and Remidez (2005), for example, reviewed the context of online debates and suggested the possible solution of training students to use the constraint-based message board as a solution to prevent the mislabeling of messages. Albeit for a different reason, Kawachi (2003) and Sorenson and Braylen (2004) likewise emphasized the importance of providing students examples of good discussion and modeling the expected behavior in order to produce better discussion and higher levels of participation.

A specific type of example is the worked example which demonstrates an expert's work to novice problem solvers (Atkinson et al., 2000). Considering that students participating in online debates are likely to be new to its instructional design, the conclusion drawn by van Gog, Paas, & van Merrienboer (2004) that learners who are novice to an instructional format use weak methods to solve problems takes on a greater degree of importance. The worked example has been found to help learners become more familiar with the format of the instruction, thus reducing cognitive load (Atkinson et al.; Li, 2005; Sweller, van Merrienboer, & Paas, 1998). The worked example has also been shown to help learners with schema construction and automation or in simpler terms, to better organize and categorize data (Li, 2005; Sweller et al.). Schema automation allows learners to become more familiar with elements of the instructional design, thus freeing working memory to think more critically about problems and less about the procedures (Sweller et al.).

With the potential for less productive learning environments, this study investigated the effects of preparing learners by exposing them to expertly-created worked examples in a hypertext format that closely emulated the behaviors expected of them in online debate. This study investigated the impact of preparing students through worked examples with a focus on student level of satisfaction, participation, and learning outcomes.

Statement of the Purpose

This study examined whether preparing learners for online debate through a worked example at the onset of the debate promoted higher learning outcomes. Students in the control group were given nothing but basic instructions about the debate procedure while students in the treatment group were given access to a worked example of the debate in addition to the instructions. The purpose of this investigation was to discover and scrutinize the effectiveness of preparing distance learners through exposure to worked examples to successfully participate in the online debate.

Consequently, its findings have broad implications for distance education. First, although the present study was offered within the context of a constraint-based online debate, its results hold the promise to inform future research into different types of worked examples and across different types of instructionally complex interactive activities. Secondly, as indicated by prior research, if worked examples were shown to offer the promise of better preparing learners, improving their perceptions, and improving their performance, instructors and course designers should, in turn, be able to create instructional ideas for promoting student interaction without the traditional limitations of apprehensive and potentially ill-prepared learners.

Research Questions

In conducting this research, the following questions were examined:

1) Did the worked example have a significant effect on students' self-reported perceptions about the online debate?
    a. Was there a difference in the perceived level of preparedness for students being exposed to the worked example and those who were not?

b. Was there a difference in the perceived level of productiveness of online debate as a learning tool for students being exposed to the worked example and those who were not?

c. Was there a difference in the perceived level of learner control for students being exposed to the worked example and those who were not?

2) Did the worked example have a significant effect on students' learning behaviors in the online debate?

a. Were there significant mean differences in the number of behaviors indicative of active participation and sustaining participation (e.g. number of postings, words per post, linguistic qualifier usage, linguistic intensifier usage, and social cue usage) for students being exposed to the worked example and those students who were not?

b. Were there differences in the distribution of cognitive skills exhibited in the postings of students being exposed to the worked example and the postings of students who were not?

c. Were there differences in the frequencies of event sequences for students being exposed to the worked example and the postings of students who were not?

Significance of the Study

Many studies have examined one or possibly more elements identified by previous literature warranting future research in regards to the effective delivery of online learning. Several studies examined the role of student perception in online learning.

These studies have come to the general conclusion that students were more satisfied with their experiences if they knew what to expect (Carswell et al., 2000; Reisetter & Boris, 2004), if they felt in control of their own learning (Clark, 2003; Scheiter & Gerjets, 2007), and if they were given the opportunity to see examples of the work they were expected to do (Kawachi, 2002; Sorenson & Baylen, 2004). Student satisfaction has been linked to more likely course completion (Conrad, 2002; Menlove & Lignugaris, 2004) and a greater likelihood that the student perceives the course to have been a more productive learning experience (Fogerson, 2005).

Looking beyond simple student perception and more directly into actual student performance, Hara, Bonk, and Angeli (2000), Henri (1992), and Schrire (2006) examined the means to code the cognitive level and critical thinking exhibited in student transcripts. While these three studies did not attempt to reach any inferentially drawn conclusions, they generated the necessary frameworks for coding the cognitive level of student messages and the critical thinking exhibited in student transcripts. These frameworks have been instrumental in supporting numerous studies that utilized transcript analysis including the present study. A pair of studies by Jeong (2003, 2005c) also looked at critical thinking by measuring event sequences within a narrower context of online debate. In these studies, Jeong developed a framework for measuring critical thinking not by coding the content of student postings, but instead by exploring the subject headings of the posts being made.

Fahy (2002), Jeong (2005b), and Jeong and Davidson-Shivers (2003) studied behaviors that sustain socially interactive behaviors finding that certain words or phrases were more likely to elicit postings from participants while other types of words and phrases were likely to do just the opposite and limit or discourage participation. Hara et al. (2000) similarly theorized that words or phrases that acknowledged other participants (i.e. social cues) were not necessarily useful for their contribution to the discussion content, but instead for their role in creating an environment that encourages and sustains the participation of others.

While acknowledging the benefits derived from online learning and other CSCL environments such as improved learning outcomes, greater higher order thinking skills, and higher levels of learner satisfaction, Joung and Keller (2004) suggested that the current research was inadequate and needed to be reinforced. Kreijns et al. (2002) identified problems in CSCL environments that could be addressed through further research into the quality of interaction: low learner satisfaction, lower quality learning outcomes, and lower participation rates. The research by Hara et al. (2000) and Henri (1992) offered several tools for examining the quality of learning outcomes in online discussion. These studies recommended analyzing online discussion through the "dimensions" of participation rate, cognitive skills, and electronic interaction patterns.

Hara et al. (2000) used the variable of raw number of postings to measure participation. Other studies identified behaviors that sustained the participation of others as also being a necessary component in ensuring sufficient participation. As a result, the present study also measured variables such as linguistic qualifiers (Fahy, 2002; Jeong, 2005b), linguistic intensifiers (Fahy, 2002; Jeong, 2005b), and social cues (Hara et al.; Vrasidas & Glass, 2000). Henri (1992) created a coding scheme consisting of five cognitive levels (later modified and used by Hara et al.) for measuring student usage of higher order thinking skills. Hara et al. measured electronic interaction patterns by coding messages as either being a direct response to another message or indirect commentary unrelated to any previous messages. The present study, however, utilized Jeong's framework for event sequence; this analysis was chosen because it was designed specifically for online debates and featured a greater number of measurable events making possible a more detailed account of the types of interaction occurring in the debate.

In utilizing the prior research, expanding upon it, and seeking a solution to recognized problems, the present study identified key indicators from the previous research in order to operationally assess the quality of the online debate (e.g. participation, cognitive behaviors, critical exchanges) as well as student perceptions (e.g. preparedness, productiveness, learner control). The current study extended prior research by examining differences in student perception and quality of learning that may be caused by preparing learners through exposure to worked examples. As a result, this research differed from the research that informed it in two major ways: (a) exploration of both student perception and learning behaviors in the same study, and (b) an evaluation framed within the context of a specific type of preparation (i.e. worked example) for a specific type of activity (i.e. online debate).

The findings of this study informed instructors about the possibility of using worked examples to prepare students to participate in an interactive constraint-based argumentation activity. The answer to the first research question identified whether exposing students to a worked example caused the student to have more positive feelings about the activity. The answer to the second question served a twofold purpose: (a) to determine if students being prepared by the worked example themselves participated more often and if students participated in ways that encouraged the participation of others, and (b) to determine if students prepared by the worked example exhibited higher levels of cognitive skill thus indicating higher order thinking and the potential for more positive learning outcomes. Generally, the cumulative effect of the answers to these questions determined if preparing students through worked examples produced higher quality interaction, higher quality learning outcomes, greater participation, and greater learner satisfaction.

Limitations

This study has the following limitations on the generalizability of its findings. The population consisted of online undergraduates enrolled in an introductory level information systems course. Seventy-seven percent (77%) of the students were older than 23, 80% were female, 90% had previously taken an online course, and two-thirds were majoring in a business-related field. Therefore, students in this study would not necessarily be representative of more traditionally-aged student populations with less experience in the online environment and a more balanced female to male student ratio.

Second, the online debate with its constraints and rigid structure creates a learning environment different from many other instructional formats. The worked example was a previous debate conducted by learners trained and given feedback to participate, selected by an instructor with vast instructional and research experience, and analyzed to ensure that it effectively modeled the desired behaviors. Consequently, the findings of this study may be difficult to replicate if researchers choose less scrutinized worked examples or evaluate the effects of worked examples on unstructured instructional formats as opposed to more structured formats like the online debate.

Definition of Terms

The following terms have been defined for use in this study:

1. Argumentation. The term "argumentation" is defined as process of developing arguments in support of a stated position as well as presenting evidence in support of the position or countering evidence opposing the position (Jeong & Joung, 2007).
2. CSCL. The term "CSCL" is an abbreviation for computer supported collaborative learning and was defined by Resta and LaFerriere (2007) as a "range of situations in which interactions take place among students using computer networks to enhance the learning environment" (p. 67).
3. Content analysis. The term "content analysis" is defined as the drawing of meaning from text by Krippendorff (2004) and the quantitative measuring of message characteristics by Neuendorf (2002).
4. Event sequence analysis. The term "event sequence analysis" is a technique identified by Jeong (2003) as the study of the relationship between messages useful for studying student interactions. He further indicated that it is aided by the hierarchical organization of online discussions allowing for each message and response to be treated as the unit of analysis.
5. Learner navigation control. The term "learner navigation control" is defined as the extent to which the learner controls the navigation of his or her learning. Higher levels of learner navigation control are thought to allow students to exercise greater self-control over their learning (Clark, 2003).

6. Linguistic content analysis. The term “linguistic content analysis” is defined as the study of human speech’s structure and nature. Interest in these types of analyses has been increased by computer advances (Neuendorf, 2002).
7. Linguistic intensifier. The term “linguistic intensifier” is defined as a word or group of words that add emphasis. Examples include words such as “very” or “extremely”. These words typically limit discussion in an online discussion (Fahy, 2002).
8. Linguistic qualifier. The term “linguistic qualifier” is defined as a word or group of words that tend to sustain online discussion. Examples include words such as “I think”, “probably”, or “if” (Fahy, 2002).
9. Navigational disorientation. The term “navigational disorientation” is defined by Scheiter and Gerjets (2007) as “a lack of knowledge concerning the structure of the hypermedia system, its extensions, and ways of accessing information” (p. 290) and this disorientation is associated with lower learning outcomes.
10. Online debate. The term “online debate” is defined as an online discussion in which individuals or small groups of students organized to argue opposing sides of an issue on the course discussion board (Jeong, 2004). In these debates, structure is added through specific rules and protocols for posting messages to the online discussion board.
11. Social cue. The term “social cue” is a cue transferred through by the user (i.e. a word, phrase, behavior) to establish a social presence in online discussion (Hara et al., 2000; Vrasidas & Glass, 2002)

12. Worked example. The term "worked example" was defined by Atkinson et al. (2000) as "an expert's problem-solving model for a learner to study and emulate" (p. 182).

CHAPTER II

REVIEW OF LITERATURE

The purpose of the chapter is to review the research literature pertinent to this investigation. In order to create a concrete foundation for this investigation, the review first evaluated the historical themes of distance education research. Secondly, it emphasized issues related to interactivity including its perceived benefits, its integration into the distance classroom, and ways to ensure its inclusion into the online environment. The review concludes by examining the importance of preparing the learner for participating in the course.

The first section of research being reviewed provides an overview of distance learning including its evolution from correspondence and the history of its research. The second section describes the use of computers to supplement the learning experience and the importance of quality interaction in an online course as well as an overview of the theories of structure, dialogue, and transactional distance. The third section examines the inclusion of argumentation, its structuring effect on dialog, and the need for scaffolding online discussion to produce quality interaction. The fourth section examines the means of and reasons for analyzing the contents of a discussion transcript including a description of how to operationally assess discussion quality. The fifth section explores the anxieties shared by online learners at the onset of a distance course and the significance of providing them adequate preparation to put to rest said concerns. Finally, the literature review concludes with a brief summary that synthesizes the findings.

Historical Themes in Distance Education Research

Research on distance education has historically attempted to gauge the quality of distance education rather than seek out actions and practices to improve it. Lockee, Burton, and Cross (1999) stated that “since the adoption of modern media for instructional purposes, innumerable attempts have been made to measure the effect that a given technology has on student achievement” (History section, ¶ 1). The history of such distance education research can, in fact, be traced back nearly eight decades when a doctoral student in his dissertation found no difference between the test scores of on-campus and correspondence students in Oklahoma (Russell, 1999). And so began a nearly endless stream of comparison studies in distance education. Lockee et al. reported the pace of these studies quickened as researchers rushed to prove that new electronic technologies such as television and radio made a positive impact on learning. They further attributed the increase in the number of studies to anxious administrators looking for positive evaluations of these new technologies. Additionally, the expense of implementing these technologies placed a burden on not only proving them as effective as traditional instruction, but better than face-to-face instruction (Meyer, 2002).

Typically, these comparison studies have come in one of two varieties. They either made a general comparison of distance learning and traditional on-campus instruction or they compared one or more types of media through which distance learning instruction is disseminated and traditional on-campus instruction. Also common to this type of research has been an approach to the problem that focuses on student satisfaction and/or achievement (Meyer, 2002).

With very few exceptions, these comparison studies have come to the same conclusion. There is no significant difference in student achievement between instruction delivered at a distance and instruction delivered on-campus and in-person. The foremost example of the prevalence of these so-called "no significant difference" studies is Russell's (1999) landmark book, *The No Significant Difference Phenomenon*. In it, he chronicled 355 media comparison studies published between 1928 and 1999 that support the notion of "no significant difference."

Clark (1994) proposed a more pragmatic vision for the future of using technology to deliver instruction. He believed that future research in this arena should focus on issues such as cost, labor intensiveness, and the cognitive efficiency that can be derived from technology-mediated instruction as opposed to simply comparing the media. It is this vision that seems to currently be most prevalent as there has been a recent and dramatic shift in the emphasis of distance education research. No longer are comparative media studies the dominant type of research. Whether one accepts potential flaws in methodology or simply accepts findings of no significant difference as settled fact, these studies have fallen largely by the wayside over the last decade. No single topic has assumed the dominant position in distance education research, but, instead, in the place of comparison studies, has come a broad array of subjects and types of research.

Lee et al. (2004) reviewed 383 articles published in *The American Journal of Distance Education*, the *Journal of Distance Education*, *Distance Education*, and *Open Learning* from 1997 to 2002. Their examination revealed that while design-related topics were the most popular category over the time period, theory and research-related topics like culture and gender issues, learning styles, and distance education history were the fastest growing. These topics, in fact, doubled from 1998 to comprise 43% of distance education research during the last year sampled. Lee and colleagues (2004) also analyzed the articles with respect to the research method being used and discovered that qualitative case studies involving the investigation of a single person, group, program, or organization were the predominant research method used in these articles. A similar review conducted by Berge and Mrozowski (2001) of the same popular journals published between 1990 and 1999 found descriptive studies the most popular research method followed by case studies. Their review also examined the frequency of articles by category. They found the three categories receiving the most attention were design issues, learner characteristics, and strategies to increase interactivity. Their findings corresponded to McIsaac and Blocher's (1998) distance learning research which identified four similar constructs of recent interest: transactional distance, learner control, social context, and interaction.

CSCL and Interaction

According to Resta and LaFerriere (2007), the term *computer-supported collaborative learning (CSCL)* includes a "range of situations in which interactions take place among students using computer networks to enhance the learning environment" (p. 67) and its usage in research was denoted as long ago as 1989.

This section of the literature review defines CSCL and identifies trends, then focuses on how and why issues concerning interaction and structure have been addressed.

What is CSCL?

Included in the range of technologies identified by Resta and LaFerriere (2007) were tools that supported communications between students either face-to-face or at a distance with the aim of enhancing their learning processes, facilitating collective learning, or facilitating group cognition. In a 20-year review of research in the field, the authors identified three major recent trends. These trends focused on new collaborative support tools, constructivist approaches to teaching and learning processes, and the creation of learning environments with greater promise for student engagement. Albeit with somewhat of an emphasis on the virtual discussion, Barbera (2006) likewise identified recent approaches to this field of research identifying such approaches as best teaching practices, structure, cognition, sociocultural aspects, and models to analyze both social and cognitive aspects.

To fully understand CSCL, it is important to understand the differences between cooperative and collaborative learning. These terms have different meanings to different researchers depending on their purposes, goals, and perspectives (Resta & LaFerriere, 2007). Cooperative learning is better suited for highly structured tasks while collaborative learning is better for less structured tasks that allow for more flexible solutions (Joung & Keller, 2004). Another distinction between cooperative learning and collaborative learning is how responsibility is distributed among members of the group. In cooperative learning, tasks are assigned on the principles of division of labor (Resta & LaFerriere) and mutual responsibility (Joung & Keller, 2004). With collaborative learning, on the other hand, tasks are distributed on the basis of mutually engaging participants (Resta & LaFerriere) and having these participants work together for the purpose of building knowledge (Joung & Keller). Resta and LaFerriere drew one further distinction concerning the general perspective of how these types of learning are viewed. Cooperative learning is based on how interaction is structured while collaborative learning is more of a philosophy. Practically speaking, Joung and Keller theorized three differences between highly structured cooperative learning (HSCP) groups and low structured collaborative learning (LSCL) groups. They believed HSCP groups might have a tendency to make better decisions, achieve a greater improvement in critical thinking, and be more interactive than their LSCL counterparts.

More generally, the literature shows a number of benefits – both realized and potential – produced from CSCL environments. Kreijns et al. (2002) recognized two of the more obvious advantages. CSCL environments have an anywhere-anytime characteristic allowing its members to be geographically dispersed and the added practical benefit of cost efficiencies due to the need for fewer instructors and their reduction in the instructional process. In addition, they also heralded its ability to generate group learning. Joung and Keller (2004) acknowledged five CSCL benefits – academic achievement, greater higher order thinking skills, increased levels of student satisfaction, better individual and group products, and better group cognition – but stressed that the research in these areas was still "shallow".

In shoring up this research that Joung and Keller (2004) called "shallow," Kreijns et al. (2002) noted the need for further exploration of some research that shows low participation rates, low quality learning, and learner satisfaction. Based on their review of the literature, Kreijns et al. identified two major problems that may be causing lower participation rates. The first problem occurs when it is assumed that social interaction will occur simply because CSCL environments make it possible. The second problem occurs when the role of social dimension on social interaction and its role in inducing off-task interactions.

The Importance of Quality Interaction

When quality interactive interventions have been developed, many previous studies indicated the potential for social interaction to create more satisfied learners who achieve higher quality learning outcomes. As was noted earlier by Kreijns et al. (2002), the CSCL environment does not guarantee a productive learning environment. In other words, interaction for the sake of interaction is not always going to bring about positive results. Instead, the quality of the interaction must also be considered and is an important factor in improving the efficacy of CSCL environments including online courses. Studies by Song (2003) and LaPadula (2003), for instance, identified problems with student engagement that could be solved through interaction. Students surveyed by Reisetter and Boris (2004) actually placed a low value on the interactive components built into the course. While students participating in Northrup's (2002) survey indicated a clear preference for interaction, they expressed frustration about being forced to participate in too many interactive assignments. They perceived the interaction as "busy work." Therefore, the challenge is to produce a consistent level of interaction – one that cultivates learning and encourages a communal atmosphere (Muirhead, 2002).

Finding a balance and effectively assimilating interaction into an online course present a number of complex issues that must be overcome. Vrasidas and Glass (2002), for example, concluded that interaction in a course cannot be severed from the context in which the course is offered. Context, in their opinion, included such factors as institutional and departmental policies, the technologies being utilized, the course content, and even the teacher. In addition, they made the point that a lack of social presence was also problematic in establishing an interactive course environment. The authors defined social presence as "the degree to which a medium allows the user to feel socially present in a mediated situation" (p. 42) and claimed that social presence increased with the number of cues being transmitted by the medium. Given the limitations of online learning, they identified several strategies to compensate for the lack of visual and aural cues typically found in the traditional face-to-face classroom. The authors further recommended capitalization, abbreviation of messages, the use of emoticons, and instructor feedback.

Shin (2003) framed the problem of student detachment in saying that the psychological distance felt by the learners may prove even greater than the physical distance. Based on the responses to the items on a survey sent to 92 online students, LaPadula (2003) proclaimed even the most highly motivated distance students may feel a sense of disengagement or isolation from the course, the instructor, the institution, and even their classmates. Integrating interaction into online courses provides a mechanism for engaging students with the course. Song (2003) attributed interaction to greater levels of achievement and a more positive attitude towards the course. Another benefit of interaction in an online course is because it helps to cultivate an online learning community within a course. Palloff and Pratt (2003) found that the "greater the interactivity in an online course and the more attention paid to developing a sense of community, the more likely students will stick with the course until its completion"

(p. 117). More strikingly, they made the claim that the development of the online, interactive learning community may be the only way to differentiate an online course from a simple correspondence course.

Many studies have indicated a student preference for interaction in distance learning environments. Muirhead (2002), for example, asserted that "students appreciate and enjoy the learning process to a greater degree when they have the opportunity to freely share with their instructor and colleagues" (p. 31). In a survey of students enrolled in an online masters program conducted by Northrup (2002), a majority of students reported liking peer interaction as well as stating a decided importance on the development of a community of learners. The results of the survey further showed these students were most comfortable with mimicking the traditional class, but doing so in an online environment.

Structure, Dialogue, and Transactional Distance

Moore and Kearsley (2005) divided interaction into three distinct types. The first type, *learner-content*, is the interaction the student has with the content or subject matter presented for study. *Learner-instructor* interaction is the interaction that exists between the learner and instructor. It includes the assistance, the testing and evaluation, and the counsel, support, and encouragement provided to each learner. The third type of interaction proposed by Moore and Kearsley, *learner-learner interaction*, is the interaction that exists between one learner and other learners. It is this type of learning that Moore and Kearsley found to be the most stimulating and motivating.

Another popular theory described by Moore Kearsley (2005) was the theory of transactional distance. Moore and Kearsley defined transactional distance as a communications gap between teachers and learners that cannot totally be attributed to geographical distance. Moore and Kearsley described the particular importance of transactional distance to distance education:

> The transaction that we call distance education is the interplay between people who are teachers and learners, in environments that have the special characteristic of being separate from one another. It is the physical distance that leads to a communications gap, a psychological space of potential misunderstandings between the instructors and the learners that has to be bridged by special teaching techniques. (p. 224)

Moore and Kearsley (2005) identified two variables affecting transactional distance: dialogue and structure. They defined dialogue as a purposeful interaction in which all parties listen to and build on the contributions of others. Many factors were found to impact dialogue. Among them were factors like class size and delivery method. There was, for example, more dialogue in small classes as well as in online classes because of the speed and frequency of responses by the instructor and learners. The second variable, course structure, was termed as "the extent to which course components can accommodate or be responsive to each individual learner's needs" (p. 226). Combined, these variables interact to form a third constraint, learner autonomy, or the degree to which a learner has to guide his or her own learning.

Based on Moore and Kearsley's theory, multiple relationships may be formed among these variables (Gorsky & Caspi, 2005; Vrasidas & Glass, 2002):

1) As structure increases, dialogue decreases, and transactional distance increases
2) As dialogue increases, structure decreases, and transactional distance decreases.
3) Transactional distance and learner autonomy are directly proportional.

It is important to note that these hypotheses have not necessarily been proven true in all cases particularly in regards to the relationship between structure and transactional distance. Twelve university instructors in distance education programs interviewed by Kanuka, Collett, and Caswell (2002), for instance, indicated reduced transactional distance occurred with both a *high degree of structure* and a *high degree of dialogue*. Vrasidas and Glass (2002) also made the claim that elements of structure, such as required participation in and moderation of discussions and group project collaboration, increased dialogue among participants. In a study by Joung and Keller (2004), it was found that high-structured cooperative groups generate more critical event sequences than low-structured collaborative groups. Even Moore and Kearsley (2005) conceded that *quality structure* could increase interaction.

There has also been findings that run contrary to transactional distance theory in general and the research that supports it. Vrasidas and Glass (2002) made the broad assertion that the theory of transactional distance is “fundamentally flawed.” They pointed to the development of better conferencing systems lessening the influence of technology and transforming the structure-dialogue relationship. As a result, it could be argued that the postulations of the theory might not ring true as distance learning technology changes, meaning the theory will likely not pass the test of time.

Additionally, Gorsky and Caspi (2005) countered the limited research supporting transactional distance. Of the three studies they identified that wholly supported the theory, they cited severe problems with construct validity. First, measurement of dialogue was based on the frequency of dialogue without regard to qualitative aspects such as learner understanding. Secondly, the definition of transactional distance used in these studies did not match the operational definition used by Moore and Kearsley (2005). Based on these problems, they concluded that transactional distance was never a valid scientific theory and it was only accepted because of its high face validity.

Argumentation

With quality interaction being shown as a significant facet of online learning, developing activities that hold the potential for such interaction becomes increasingly important. One such activity is the purposeful introduction of argumentation to online discussion. Oh and Jonassen (2006) defined argumentation as the construction and comparison of arguments using various types of reasoning. They further noted the potential of argumentation as a problem solving activity.

In this section, the concept of argumentation will be discussed by first developing the link between structured dialog and argumentation, then exploring the use of scaffolding tools in helping students think more critically in argumentation exercises, and concluding with an examination of using constraints such as those seen in online debate to scaffold argumentation.

Structuring Dialog

It is important to note that interaction for the sake of interaction is not always going to manifest itself in a productive learning environment. In other words, increasing interaction through the introduction of online discussions and the like is not going to transform an online course into a learning environment full of cooperative learners eager to seize upon the opportunities presented by increased leaner-learner interaction. Instead, the quality of the interaction must also be considered. Although many ways to create chances for quality interaction exist, paying greater attention to the structure and organization of online discussion appears to be particularly promising. One benefit, according to multiple sources, is that structure reduces off-task talk and leads to more focus on the topic (e.g. Hirsch, Saeedi, Cornillon, & Litosseliti, 2004; Jeong & Joung, 2004; Jonassen & Redmidez, 2005). As collaborative online learning places additional demands on learners, structuring dialog also holds the potential to remove some of the cognitive load from their shoulders and better facilitate their participation (Hron & Freidrich, 2003). In contrast to the hypothesis in his transactional distance theory that increased structure limited interaction, Moore and Kearsley (2005) conceded that quality structure can, in fact, improve interaction.

Of course, human nature dictates that within an online discussion, the thoughts of one participant may sometimes come in conflict with the thoughts of other participants. If well-structured, however, this is not necessarily anything to fear. Structured dialogue creates situations ripe with opportunity for students to argue issues in a civil, organized manner thus promoting meaningful interaction. Argumentation, for example, has been heralded because it allows the learner the opportunity to express his or her thoughts and then weigh evidence in support of these thoughts (Hirsch et al., 2003; Jeong & Joung, 2004).

Scaffolding Argumentation

If argumentation can be used to bring forth higher quality interaction, then what can be done to develop these skills among online learners? This is the question asked by Cho and Jonassen in a 2002 study. Their answer was to use cognitive tools to scaffold argumentation. In their study, 69 undergraduate economic students were asked to participate in a problem solving group discussion with half doing so in a threaded discussion and the other half using a scaffolding tool. Their findings showed that students using the scaffold used more problem-solving comments in six of the seven measured categories. In addition, they found that scaffolded discussions produced more components of argumentation than did unscaffolded discussions.

In a study of 58 undergraduate teacher education students, Oh and Jonassen (2007) likewise found benefits of scaffolding online argumentation. The students assigned to the scaffolded discussion group both generated and tested a greater number of hypotheses in their postings than those students participating in threaded discussion. The researchers also cited argumentation scaffolding as assisting novice problem solvers in generating evidence to support their arguments. Based on student comments in a qualitative case study, Hodgkinson-Williams and Mostert (2005) concluded that students valued argumentative scaffolds for their "importance of reflection" and the avoidance of "impulsive (and subjective) counter arguments and interjections" (p. 102).

Constraint-Based Argumentation

Placing constraints on the messages being posted appears to be a promising way to scaffold the argumentation process. Jeong and Joung (2007), for example, identified requiring students to embed constraints within online discussion as an approach to scaffold argumentation. One way to add structure to the arguments is by placing constraints on the messages being posted. Inserting specific message labels into the subject headings has been shown to make argumentative exchanges more explicit (Jeong, 2004; Jonassen & Remidez, 2005). Discussions using label constraints have popularly been described as online debates. Despite the constraining nature of these labels, Jeong and Joung (2004) found the use of labels actually generated more posts. This paradoxical finding may be explained at least, in part, by their assertion that the labels assisted students in examining the structure of arguments in discussion threads. Another reason proposed by Jeong and Joung (2004) was that the labels allowed students to quickly locate any points of contention within the discussion. The highly structured cooperative learning environments that constraint based argumentation would facilitate have been linked to better decision making, greater levels of critical thinking, and more interaction (Joung and Keller, 2004). Beyond the assistance these labels afford current online students, they further described these labels as a practical method for obtaining data for future research.

Implementing constraint-based argumentation is not totally without obstacles to overcome. In subsequent studies, Jeong and Joung (2004, 2007) found constraints with labels inhibited many of the processes necessary for critical argumentations. They presented two possible reasons for this inhibition. First, the labels reduced the likelihood that students would challenge the messages of other students. Secondly, the labels seemed to create a tendency for students to shy away from responding back to messages critical of their previous claims. Although presented in the context of reasons for differences between the nature of critical responses made by males and females, the suggestion made by Jeong and Joung (2004) that the explicit nature of the labels "heightened the perception of critical responses as excessively confrontational" (p. 4) might also explain the lack of critical analysis in constraint-based argumentation.

Regardless of the reason, it is very clear that the lack of critical analysis is problematic as it serves to inhibit argumentation and thus interaction among students. Perhaps lending credence to Moore's transactional theory is the finding that the students sometimes struggle with the structure itself. Although students in their study were found to generally label messages correctly, Jonassen and Remidez (2005) still noted problems by students with the labeling of messages. For example, students struggled with correctly identifying messages as "personal opinion or belief." In addition, they observed instances of compound messages. In these messages, students would offer an initial point and then modify or qualify their claim within the same message. As a result, Jonassen and Remidez suggested "that students need some formal introduction and perhaps training on how to use a constraint-based discussion board intended to support argumentation before they will be able to fully utilize these types of tools" (p. 127).

Assessing Discussion Quality

With quality interaction being cited as key component in the efficacy of online discussion, gauging the quality of interaction and dialogue takes on a vitally important role in the research. One way to do this challenge is to analyze the content of the discussions themselves.

In this section, the field of content analysis will be discussed generally as well as how content analyses are interpreted through the development of coding protocols and the behaviors and constructs measured through content analysis for the purpose of measuring critical thinking.

Content Analysis

Krippendorff (2004) defined content analysis as "a research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use" (p. 18). More succinctly, Neuendorf (2002) called it "the systematic, objective, quantitative analysis of message characteristics" (p. 1). Gunawardena et al. (1998) summarized the need for content or transcript analysis, "To settle for such measures [online or paper surveys] in evaluating computer conferences is to overlook the unparalleled opportunity to observe knowledge construction in progress offered by transcript analysis. Transcripts give us participants' own statements, specific message, or group of messages" (p. 2).

Content analysis can be classified in a number of different ways. One of the more fundamental ways to categorize content analyses is the type of document being analyzed such as scholarly journals, literary works, newspapers and periodicals, or of a more recent vintage the online discussion transcript (Neuendorf, 2002). Perhaps an even more elemental way to break down content analysis is to look at what unit is being analyzed. The unit of analysis can be based on themes, messages, or more precisely, the individual sentence (de Wever et al., 2006; Fahy, 2001; Neuendorf). One more way to categorize content analyses according to Neuendorf is through the context in which the analysis is framed. One example is thematic content analysis which scores messages for the purpose of evaluating the psychological characteristics of the person. Another context is stylometrics, or the study of the style of language being used. And yet another is the linguistic content analysis which is the study of human speech and its structure.

Historically, the first content analysis is recognized to have occurred in Sweden during the 1700s (Krippendorff, 2004). However, it is the much more recent introduction of computer technology and its impact on content analysis that better informs this discussion. Beginning in the 1950s, the computer began to be put into use for the purpose of content analysis (Krippendorff). The progression of computer-aided analysis has continued largely unabated due to its ability to process large amounts of data with speed and precision (Krippendorff; Neuendorf, 2002). As evidence, today there exist numerous quantitative computer text analysis programs. These programs vary widely in functionality. Some produce alphabetical listings of word counts. Others compare texts to built-in dictionaries for the purposes of placing words or phrases into categories and then producing category frequencies. Still others perform more complicated analyses such as finding key words and their context in the document.

Despite its promise, the field of content analysis (and within it computer-aided text analysis software) has drawn a degree of criticism. Krippendorff (2004), for example, while generally supportive of computer-aided text analysis software does call the computer’s ability to recognize only strings of characters and the corresponding inability to understand abstraction a hitch. There are a number of smaller problems with analyzing the content of transcripts -- most of which are of a more practical nature. For instance, there is no guarantee that the text transcript will be presented in a format compatible with the software (Rourke et al., 2001). Fahy (2001) attributed problems with previous transcript analysis studies to a pair of causes -- the complexity of the coding instrument resulting from too many categories or ambiguities between these categories and the focus on the wrong unit of analysis namely in the fact that these studies used something other than the sentence.

Development of a Coding Protocol

When performing a quantitative content analysis, one of the key tasks is the development of a coding protocol or scheme. In referring to the development of a coding protocol, Rourke and Anderson (2004) wrote that "such an instrument should be sensitive to instructional interventions such as training students in the problem-solving process or providing expert assistance while students are engaged in problem-solving tasks" (p. 14). The development of coding schemes, in fact, accounted for most of the published quantitative content analysis research according to Rourke and Anderson. Informed largely by a literature review performed by Rourke and colleagues (2001), Rourke and Anderson reviewed procedures to create a "theoretically valid protocol" and recommended five steps (p. 8): (a) identifying the purpose of the coding data, (b) identifying behaviors that represent the construct, (c) reviewing the categories and indicators, (d) holding preliminary tryouts, and (e) interpretation of the coding scheme. It is largely this fifth step that guides this section of the literature review.

One of the first decisions for researchers to make when deciding how to interpret the results of a content analysis is to determine the units they will be analyzing. Perhaps an even more elemental way to break down content analysis is to look at what unit is being analyzed. The unit of analysis can be based on themes, messages, or sentences (de Wever et al., 2006; Fahy, 2001; Neuendorf, 2002). Gunawardena et al. (1998) explained that "a degree of subjectivity in doing this type of analysis" is unavoidable because researchers are "clearly influenced by their own conceptual frameworks and cultural knowledge" (p. 4). It is this lack of a precise definition for unit of analysis leads to confusion (Schrire, 2006). Efforts by de Wever et al., Rourke et al. (2001), and Rourke and Anderson (2004) attempted to lend clarity by evaluating the relative merits of the different types of analysis units by examining previously conducted content analysis studies. Typically, these studies featured discussion in which the researchers justified their selection of a particular analysis unit.

In describing their study, Hara et al. (2000), for example, wrote they utilized the paragraph unit of analysis because "it was assumed that each paragraph in a submission was a new idea unit since college-level students should be able to break down the messages into paragraphs" (p. 122). Rourke et al. (2001), however, called this conclusion "optimistic" and considered the paragraph unit of analysis meaningless due to the inability of college students to do so. De Wever et al. (2006) and Rourke et al. argued instead that coding schemes should use the message as the unit of analysis because it is defined by the author thus making their identification objective. As for the viability of the message unit of analysis, de Wever et al. evaluated fourteen content analysis schemes and found fifty percent of them used the message as the unit of analysis.

Upon determining the elemental unit of measurement, scoring these units and scoring them reliably becomes important. The first step in this process is the training of coders (de Wever, 2006; Rourke et al., 2001). Neuendorf (2002) offered three words in describing the steps in preparing coders: “Train, train, and train” (p. 133). She described a series of steps necessary to train coders. One of the first steps is the discussion of the measured constructs with the coders as well as a practicing of the coding scheme with the coder. The coders should then practice the code independently and discuss their results with the trainers. The process generally concludes with the coders performing a pilot coding of a subsample of the data for the purpose of ensuring the reliability of the coding scheme. The end result of following a series of well-defined steps is the development of well-trained coders. Rourke et al. linked well-trained coders with increased interrater reliability.

Rourke et al. (2001) defined interrater reliability as “the extent to which different coders, each coding the same content, come to the same coding decisions” (p. 6) and stated that it led “ultimately to replicability (the ability of multiple and distinct groups of researchers to apply a coding scheme reliably)” (p. 7). Several statistics exist for measuring interrater reliability (Table 2.1). De Wever et al.’s (2006) analysis of previously performed content analyses indicated that, when reported, the percent agreement and Cohen’s kappa statistics were the most common methods of reporting interrater reliability.

Table 2.1

Measurements of Interrater Reliability

Measurement		
Percent Agreement	Formula	$PA = F_A^* / n$
	Advantages	Simple to administer (de Wever et al., 2005; Neuendorf, 2002; Rourke et al., 2001)
	Disadvantages	Does not account for agreement by chance (de Wever et al., 2005; Neuendorf, 2002; Rourke et al., 2001) Coded scores must match exactly (de Wever et al., 2005; Neuendorf, 2002)
Cohen's kappa	Formula	$k = (F_A - F_C) / (n - F_C)$
	Advantages	Calculation takes into account chance agreement between coders (de Wever et al., 2005; Rourke et al., 2001)
	Disadvantages	Statistic is overly conservative (Neuendorf, 2002; Rourke et al., 2001) No specified level of acceptable agreement (Rourke et al., 2001)
Krippendorff's alpha	Formula	$\alpha = 1 - (Do - De)$
	Advantages	Calculation takes into account chance agreement between coders (de Wever et al., 2005; Neuendorf, 2002; Rourke et al., 2001) Takes into account the magnitude of misses (de Wever et al., 2005; Neuendorf, 2002) Adjusts for nominal, ordinal, interval, or ratio data (de Wever et al., 2005; Neuendorf, 2002; Rourke et al., 2001)
	Disadvantages	Difficult to calculate (Neuendorf, 2002) Statistic is overly conservative (de Wever et al., 2005)

* F_A = number of agreements between coders; F_C = number of agreements between coders expected by chance; n = total number of units coded for the test

Despite the popularity of the Cohen's kappa statistic, de Wever et al. (2006) recommended Krippendorf's alpha as a more robust statistic due to its ability to account for the magnitude of misses. It was those studies, however, that failed to report any interrater reliability that proved to be particularly problematic. Only 9 of the 14 in the de Wever et al. study and 10 of the 19 in the Rourke et al. (2001) study reported interrater reliability data. Rourke et al. suggested that failure to report interrater reliability rendered a study virtually useless while de Wever et al. cited this failure as holding back research in this area.

Measurable Behaviors and Constructs

In addition to the research about how content analyses were performed and coding schemes were developed, a number of studies have looked at what behaviors and constructs are being measured. Henri's 1992 study seems to be the basis for most of this research (Rourke and Anderson, 2004). All of the literature that informed this section of the review cited this work. Henri's goal was to "identify the elements within messages which would tell us something about the ways people learn" (p. 129). One way that many researchers attempted to reach this goal is by studying cognitive potential by measuring higher-order and critical thinking skills (e.g. de Wever et al., 2006; Hara et al., 2000; Henri; Marra, Moore, & Klimczak, 2004; Rourke and Anderson; Schrire, 2006).

Henri (1992) developed her cognitive skills coding scheme (Table 2.2) with the expectation that clarification and inference activities within an online discussion were indicative of knowledge acquisition while if only a superficial processing of information occurs during a discussion, it was indicative of three less desirable learning outcomes: problems with the task at hand, lack of knowledge, or a lack of in-depth processing. To measure these skills, Henri developed a framework consisting of five levels: elementary clarification, in-depth clarification, inferencing judgment, and application of strategies.

Table 2.2

Henri's Cognitive Skills Framework

Reasoning Skills	Definition
Elementary clarification	Observing or studying a problem, identifying its elements, and observing their linkages in order to come to a basic understanding.
In-depth clarification	Analyzing and understanding a problem to come to an understanding which sheds light on the values, beliefs, and assumptions which underlie the statement of the problem.
Inferencing	Induction and deduction, admitting or proposing an idea on the basis of its link with proportions already admitted as true.
Judgment	Making decisions, statements, appreciations, evaluations and criticisms.
Application of strategies	Proposing coordinated actions for the application of a solution, or following through on a choice or a decision.

Note: Hara et al., 2000 (p. 15)

Because of the difficulty in operationalizing critical thinking and higher-level learning outcomes, Schrire (2006) concluded that most frameworks were based on one or more taxonomies. One of the more often used taxonomies for this purpose is the cognitive domain of Benjamin Bloom's 1956 Taxonomy of Educational Objectives. The cognitive domain for Bloom's Taxonomy described six levels (Table 2.3) of cognitive activity -- each level representing a higher order process than the last. This process ranged from recall or facts to the further development of intellectual skills and abilities (Reigeluth & Moore, 1999). The first three levels were thought to be foundational for the remaining three levels which were thought to represent more complex cognitive activity (Reigeluth & Moore). Bloom's taxonomy has proven a popular choice as a framework to assess higher-order learning outcomes because its categories allowed researchers to easily map learning activities to learning outcomes (King & Duke-Willams, 2001).

Table 2.3

Revised Bloom's Taxonomy

Level	Description
Knowledge	Students working at this level can remember and recall information ranging from concrete to abstract.
Comprehension	Students are able to understand and make use of something being communicated. In this level, student can translate, interpret, and extrapolate the communication.
Application	Student can apply appropriate concepts or abstractions to a problem or situation even when not prompted to do so.
Analysis	Students can break down the material into its parts and define the relationship between the parts.
Synthesis	Students create a product, combining parts from previous experience and new material to create a whole.
Evaluation	Students make judgments about the value of materials, ideas, and so forth.

Note: Adapted from Reigeluth & Moore, 1999

Although Henri (1992) made no acknowledgement of the similarity between the levels of her analytical framework and the levels of Bloom's taxonomy, Hara et al. (2000) acknowledged similarities between Henri and Bloom. They likened Henri's "elementary clarification" level to Bloom's knowledge level, her "in-depth clarification" level to the comprehension level, the "application of strategies" level to the application level, the "inference" level to the synthesis level, and the "judgment" level to the evaluation level. In the subsequent content analysis on cognition performed by Hara et al., they utilized Henri's framework. In evaluating the content of online discussion, Schrire (2006) also categorized messages on the basis of Bloom's taxonomy. Her scheme was more simplistic, however, by simply categorizing messages falling in the lower three levels as "lower-order thinking" and those messages falling in the higher three levels were categorized as "higher-order thinking".

In addition to exploring the cognitive skills of the students, many researchers have identified the need to also explore the social dimension of online discussion. The reasons for examining this dimension are varied. For example, Henri (1992) described the underlying practical implications of social activity writing that it was important for "participation, social cohesion within the group, and the feeling of belonging" (p. 126). Henri further described the importance of the social element from the perspective claiming that it helped produce a greater efficiency of message exchange. Reasons given for this greater efficiency were more information within the group environment, a better circulation of ideas, and the establishment of links among the participants. De Wever et al. (2006) suggested that researchers of cognitive constructivism believed that social transaction in CSCL environments led to knowledge elements being made more explicit and the consecutive reorganization of these elements. Wickersham and Dooley (2006) linked deep learning with active engagement and the development of cognitive skills with a social context. Schrire's (2006) findings likewise supported the social construction of knowledge.

As is the case with the cognitive dimension, there are a number of methods and statistics to measure the social dimension of online discussion. Some of the statistics are used to simply measure level or amount of participation. Examples include number of postings (e.g. Herring, 1993; Jeong & Davidson-Shivers, 2003) and length of postings (e.g. Barrett & Lally, 1999; Herring, 2000). Others are more latent measures designed to measure social behaviors. A high number of linguistic qualifiers or a low number of linguistic intensifiers, for example, would be indicative of such behaviors. Linguistic qualifiers are words or phrases that tend to encourage or sustain future dialogue. Qualifiers singled out by Fahy (2002) and Jeong (2005b) were "but", "if", "may/might", "I think", and "often". To this list, Fahy (2002) also identified "probably" and "though" as qualifiers. Linguistic intensifiers, on the other hand, are words or phrases that tend to limit future dialogue. The five most popular intensifiers identified in studies by both Fahy (2002) and Jeong (2005b) were "very", "only", "every", "never", and "always".

Rummel, Spada, and Hauser (2009) similarly placed a high value on a set of behaviors they called *sustaining mutual understanding* that were tailored to contribute to their colleagues' knowledge. Rummel et al. also heralded a set of behaviors they called *reciprocal interaction* that helped participants contribute in equal measure. Another socially redeeming behavior within online discussion would be the use of social cues. Social cues are words, phrases, or behaviors not related to the content that acknowledge others (Hara et al., 2000).

There are other issues related not to how previous studies have been conducted but instead on what these studies fail to measure. The electronic interaction pattern is the third type of measurement needed to thoroughly operationally assess online discussion. While Jeong (2003) heralded content analysis for its ability to focus on "on the quality of messages in relation to performance in critical thinking and argumentation," he also believed that previous studies of this nature "fall short in providing a robust methodology for measuring student interactions and examining how specific event sequences affect subsequent discussion and cognitive outcomes" (p. 26). According to Jeong (2003), the very group interactions that content analysis failed to measure are the very thing that supports critical thinking and the creation of new knowledge. As a supplement to traditional content analysis, he proposed that the sequence of messages be analyzed. In a later research presentation, Jeong (2005c) called event sequence analysis the "missing factor" in the research on interaction.

To measure these sequences, Jeong (2005a) developed a spreadsheet macro entitled the *Discussion Analysis Tool* (DAT) to count the frequencies in which type of response elicited another type of response. This particular tool is well-suited to constraint-based argumentation because the constraints allow the message types to be entered very quickly and correctly into DAT. In describing the usefulness of DAT, Jeong (2003) said that "tools such as DAT will be useful for empirically testing interactions and structures that enhance online discussions, providing the basis for more systematic testing of instructional interventions and computer-conferencing technologies" (p. 25). As an example of the types of conclusions that could be drawn from DAT, Jeong (2003) found that interaction featuring conflicting viewpoints produced more discussion and a greater level of critical thinking.

Preparing Learners

There appear to be many barriers that tend to constrain quality interaction -- most notable among them are those factors that limit student participation. Feelings of isolation and detachment are but one factor. Yet another is the lack of frequent, substantive feedback from the instructor (Blignaut & Trollip, 2003). A third barrier identified by Vrasidas and Glass (2002) was that using the computer required learning new technological and discourse skills that further hampered effective participation in online environments.

This section addresses the need to prepare learners for the rigors of online debate and other interactive online activities, the role of scaffolding, the potential of worked examples as a means of preparedness, and the application of a worked example from a recent study.

The Need for Preparedness

Previous research indicated the most pressing problem constraining quality interaction concerns the clarity of what is expected of the learners. Fogerson (2005) described student readiness as a prerequisite for a satisfactory and effective learning. Carswell et al. (2000) concluded that role confusion and unclear expectations contributed to a lack of participation by students. In another study, 95% of distance learners stated "explicit expectations" were important and 91% cited a need for clear course procedures (Reisetter & Boris, 2004). The need for instructors to provide clear expectations has been echoed in a number of other studies (Bozarth, Chapman, & LaMonica, 2004; de Bruyn, 2004; Northrup, 2002; Palloff & Pratt, 2003). In a discussion-laden course, the need for clearly stated expectations grows only greater as learners have reported increased levels of dissatisfaction with the inclusion of additional channels of information (Reissetter & Boris).

When constraint-based online debates are added to online courses with their emphasis on structure, hypertext navigation, message labels, and the like, the need for clearly stated expectations takes on even greater significance. The challenge for distance learning instructors becomes how to relay clear expectations to learners particularly in a complex debate setting. Kawachi (2003) offered that "distance students want or would like to see, at the onset of the task, samples of an assignment to gauge the product quality required" (p. 76). Modeling the expected learning behavior is an essential learning characteristic of scaffolding (Sam, 2005). Sorenson and Baylen (2004) likewise emphasized the importance of providing students with examples and criteria of good discussion. They argued that if students understood not only their roles, but also how to post and use the discussion board, higher levels of participation should ensue.

Two of the three elements listed in Sam's (2005) scaffolding model apply directly to the idea that students need preparation for online discussion. The first element, *content scaffolding*, refers to the guidance provided to students in order for them to learn how to perform a task. The second element, *procedural scaffolding*, refers to the guidance provided to a student in order to use resources to learn how to perform a task. Sam stated that procedural scaffolding was just as important as content scaffolding. Wong-Bushby et al. (2005) examined the influence of process and content scaffolding on online discussion. Their findings indicated that the integration of a process scaffold alone into an online discussion did not increase group satisfaction, but the integration of content and process scaffolds together increased both satisfaction and learning. This finding is partially affirmed by Rummel et al. (2009) where learners were provided both a content (i.e. example) and process (i.e. elaborative support by way of instructor prompting) scaffold before engaging in a computer-mediated, collaborative activity. Participants exposed to the both types of scaffolding reported the scaffolds to be more useful and facilitate knowledge transfer into the activity than those exposed to different types of scaffold.

Based on a review of previously conducted studies, Scheiter and Gerjets (2007) presented another compelling reason for preparing the learner to participate in debate. The authors cited a "navigational disorientation" that results from "a lack of knowledge concerning the structure of the hypermedia system, its extensions, and ways of accessing information" (p. 290) and this disorientation is associated with lower learning outcomes. Chen (2003) offered a similar definition stating that navigational disorientation stems from students having "trouble following the line of discussion because messages do not flow in a logical order" (p. 25) and they do not know how to move from location to location. When learners better understand the navigational aspects of a hypertext structure, they are said to have greater learner control. Clark (2003) found higher levels of learner navigation control allowed students to exercise self-control over their own learning.

Consequently, preparing the learner to work in this environment might build a level of procedural knowledge to lessen the effects of disorientation and increase learner control. In addition, according to Hirsch et al. (2003), a visual preparation presents a possible twofold benefit of visually presenting the debate example: reduction of cognitive load and assistance in message exchange. Pedersen and Liu (2002) concluded that non-linear access to information facilitates learning without necessarily dictating the sequence of the learning and was good for users both because it was under their control and because it allowed for numerous viewings. In their subsequent study, they offered learners the opportunity to see an expertly-crafted hypermedia presentation and found that learners emulated the expert behaviors being monitored.

Worked Examples

Specific among the techniques to better prepare learners for successful participation in online debates are exposing them to worked examples. Worked examples provide "an expert's problem-solving model for the learner to study and emulate" and have received considerable amounts of research attention over the past few years (Atkinson et al., 2000, p. 182). From this research, a number of worked example advantages have been identified. These advantages range from skill acquisition and transfer performance (Atkinson et al.; Li, 2005; van Gog et al., 2004) to problem solving (Atkinson et al.; Li; Sweller et al., 1998; van Gog et al.) to two areas of particular interest: schema construction and the reduction of extraneous cognitive load.

Worked examples facilitate improved schema construction (Li, 2005; Sweller et al., 1998). Sweller et al. defined schemas for their ability to organize and categorize information based on how the user will use it. Further, these schemas are stored in long-term memory as opposed to working memory. In answering why this particular difference is critical, Sweller et al. made an analogy between how expert and novice chess players store and process board configurations:

> Why should memory of board configurations result in superior playing skill? Skilled chess players recognize most of the board configurations they encounter, and they have learned the basic move associated with each configuration. Unlike

> less-skilled players, they do not have to search for good moves using limited working memory … All studies confirmed that the major factor distinguishing novice from expert problem solvers was not knowledge of sophisticated, general problem-solving strategies but, rather, knowledge of an enormous number of problem states and their associated moves. (p. 254)

At the core of schema construction is schema automation. By automating schema construction, familiar tasks are performed from long-term memory while unfamiliar tasks are performed from working memory (Sweller et al.). Consequently, the practical application of this theory is that learners can use working memory to search for solutions to problems and to think critically about these problems.

Another oft-cited benefit of worked examples are their ability to reduce extraneous cognitive load (Atkinson et al., 2000; Li, 2005; Sweller et al., 1998). Extraneous cognitive load was defined by Sweller et al. as the unnecessary cognitive load placed on the learner by the format of the instruction. A possible reason is that learners who are novice to an instructional format employ weak methods to solve whatever problem is presented thus hindering learning (van Gog et al., 2004). By developing instructional interventions such as worked examples, Sweller et al. stated that learners would become more familiar with the instructional tasks thus reducing extraneous cognitive load and improving learning.

Sweller et al. (1998), however, cautioned that the worked example should be presented in an integrated fashion. The authors claimed that worked examples that required users to look in two or more places for information actually led to a phenomenon known as split-attention that actually led to an increase in extraneous cognitive load. The authors cited research that indicated learners performed better on computer applications without a manual because the manual represented a second channel of information. However, if one source of information cannot be understood in isolation from other sources of information, the authors believed that worked examples consisting of multiple sources could reduce extraneous cognitive load given that the sources were not redundant and were physically integrated.

Application of a Worked Example

A realistic application of the worked example research can be seen in Rummel et al. (2009). In this research, the authors had previously established the hypothesis that observation of a “worked-out collaboration example” that demonstrated elements consistent with good collaboration would allow students to learn more about good collaboration and thus improve their performance within an unsupported collaborative environment. The authors found the worked example modeled three key behaviors: coordination of collaborative activity, time management, and knowledge transfer into the processes of problem solving.

In their current study, Rummel et al. (2009) expanded upon this idea to see what type of example would be most effective in preparing learners in a complex, computer-mediated collaborative setting. The research sought to determine if a worked-example (i.e. model condition), script condition, or a combination of the support with instructional prompting worked better in preparing learners to work in a collaborative environment. Rummel et al. defined the collaborative scripts as “promot[ing] a fruitfully structured interaction by giving precise instructions on how to interact, thus improving the joint problem solving and knowledge acquisition" (p. 73).

Procedurally, the students were to undergo a two-person collaborative problem-solving exercise. Before the exercise, they were assigned to five groups -- one group served as a control and four groups were allowed to observe a different experimental model, script, or hybrid condition before moving on without further support into the problem-solving exercise. Their interaction was coded on nine different measures using a five-point scale ranging from very bad to very good and each student completed a posttest and questionnaire to gauge their perceptions about the experience.

Rummel et al.'s (2009) conclusion was that the students being exposed to the worked example exhibited better collaboration than seen in the control group (who were exposed to no additional support) and the group exposed to the collaborative scripts. The students in the model condition group better managed their time and dialog, better divided tasks, and maintained a higher task orientation. When evaluating student perceptions, the results were not as clear. The students in the model condition group generally responded with more favorable attitudes about their interest in the activity, the helpfulness of the worked example, and ability to transfer skills learned from the model condition to the exercise than those in the script condition group, but their responses were not significantly more positive than those in the control group.

Summary

The history of distance learning spans from the meager delivery of course material via mail to today's asynchronous online course offered over high-speed Internet.

Throughout much of its history, distance learning researchers have attempted to prove its equivalency to the traditional face-to-face course as opposed to exploring ways to improve the quality of instruction within distance education. Only recently has the research emphasis shifted away from comparison studies to issues more closely related to instructional quality.

Among an array of topics at the forefront of today's distance learning research is the integration of interaction into the asynchronous online course. Multiple sources have linked interaction to greater levels of student achievement (Song, 2003), more positive attitudes (Muirhead, 2002; Palloff & Pratt, 2003; Song), and higher retention ratios (Palloff & Pratt) as well as those studies that show a student preference for interaction. Despite their potential, Kreijns et al. (2002) indicated less positive outcomes for online learning and other CSCL-based environments such as low participation, low quality learning, and low learner satisfaction. As a possible explanation, Kreijns et al. proposed that these less positive outcomes were due to instructors assuming social interaction would occur simply because of the technology. Consequently, it appears that the interaction must be carefully considered before being introduced into the learning environment.

One way to produce quality interaction is through the introduction of structure. A pair of studies by Jonassen and colleagues (Cho & Jonassen, 2002; Oh & Jonassen, 2007) demonstrated students generate more hypotheses and solve more problems when using scaffolding tools. Many studies have reported structure to improve interaction within online environments by focusing discourse (e.g. Hirsch et al., 2004; Jeong & Joung, 2004; Jonassen & Redmidez, 2005) and reducing the cognitive load inherent in many online courses (Hron & Freidrich, 2003). One way to add structure identified in the previous research is by placing constraints on the messages being posted. Constraining messages through the insertion of specific labels into the message headings being posted has been shown to make exchanges more explicit (Jeong, 2005c; Jonassen & Remidez) and generate a greater number of postings (Jeong & Joung, 2004). Due to the collaborative exchanges resulting from the insertion of these labels, this type of constraint-based argumentation commonly assumes the title of an online debate.

When combined into an online debate, previous research indicates increased structure, argumentation constraints, and the hypermedia environments together can create a number challenges. Jeong and Joung (2004) believed that the labeling schemes prevalent in online debate led to a lack of critical responses. Jonassen and Remidez (2005) noted that students struggled with the structure of the labeling scheme itself. Scheiter and Gerjets (2007) discussed a "navigational disorientation" that results from interacting with hypermedia environments.

Previous research also presented a possible solution to these problems. That solution comes in the way of learner preparedness. Jonassen and Remidez (2005) as well as Scheiter and Gerjets (2007) both asserted that learner preparedness might be the answer to the problems presented by structure and navigational disorientation. More generally, much of the research also indicated the need for treatments to better prepare all students for the online discussion environment. As a result, instructors need to provide students clearly stated expectations at the beginning of the course (e.g. Bozarth et al., 2004; Carswell et al., 2000; de Bruyn, 2004; Northrup, 2002; Palloff & Pratt, 2003; Reissetter and Boris, 2004).

In reviewing different ways to prepare students to participate in online education, worked examples offer a number of advantages particularly in the cognitive realm. Atkinson et al. (2000) defined worked examples as a model developed by experts and studied by novice learners for the purpose of preparation and emulation. Previous research suggested cognitive advantages presented by worked examples including skill acquisition and transfer performance (Atkinson et al.; Li, 2005; van Gog et al., 2004), problem solving (Atkinson et al.; Li; Sweller et al., 1998; van Gog et al.), schema construction (Li; Sweller et al.), and cognitive load reduction (Atkinson et al.; Li; Sweller et al.).

A great deal of the current research is geared towards looking at how students learn and doing it by analyzing the content of the student's own discussion postings. In large part, this rise in content analysis studies is enabled both by the ready availability of discussion transcripts (Neuendorf, 2002) and the advancements in quantitative content analysis software (Krippendorff, 2004; Neuendorf). Research indicated a range of behaviors that can be measured through content analysis. Social elements, for instance, have been measured both because they have been found to increase participation (Fahy, 2002; Henri, 1992; Jeong, 2005), but also because they have been shown that social activity allowed students to cognitively construct knowledge (de Wever et al., 2006; Schrire, 2006; Wickersham & Dooley, 2006). Higher-level and critical thinking skills have been measured through complex coding protocols in many studies (Hara et al., 2000; Henri; Rourke & Anderson, 2004; Schrire). In many instances (e.g.. Hara et al.; Henri; Schrire), these protocols have been based on the cognitive domain of Bloom's

Taxonomy with discussion units being categorized as either higher-order or lower-order by coders depending on where they deemed to fall within the taxonomy.

CHAPTER III

METHODOLOGY

The review of literature indicated there is a need to study ways to better prepare learners for complex interactive activities. The research suggested that not only does the quality of the behaviors demonstrated in the activity need to be investigated, but also that the perceptions learners bring into these activities need to be examined. The research further suggested that expertly-crafted worked examples hold the potential to better prepare online learners. This chapter describes the research and analysis methodologies used to investigate the effects of preparing learners through a worked example and is divided into six sections: (a) a description of the student population, (b) an examination of the online debate process, (c) a statistical analysis of the worked example provided to the students in the treatment group, (d) a description of the student perception instrument, (e) a report of how data was collected, and (f) a description of the data analysis techniques employed.

Setting

The online course used for this study was offered asynchronously using the WebCT course management system. The course was one semester in duration and taught entirely online. All coursework and course participation were conducted via WebCT.

The material presented in the course spans from basic computer jargon to emerging trends in information systems with an emphasis on their relationship to business processes. Some examples of these topics included an overview of terminology related to computer hardware, the importance of databases to better market goods and services to customers, the growth of electronic commerce, and issues relevant to information privacy and ethics.

Description of the Population

The participants in this study were undergraduate students from a small, liberal arts university in the Southeast region of the United States enrolled in a semester-long introductory level information systems course. The study was conducted over the course of two of these semesters. The course was a required course for students majoring in management information systems, a major elective for students majoring in one of the other business disciplines, and a computing skills elective for numerous other majors on campus. As a result, students enrolled in this course represented a wide cross-section of majors across campus.

Online surveys were used to collect data from students. Among the data being collected were demographic information about the participants. Sixty-one students participated in the surveys. The sections to follow described the characteristics of the population.

Age, Gender, and Ethnicity of Participants

The data presented in Table 3.1 identifies the ages of the participants distributed across four categories. The plurality of the participants (34%) were between the ages of 24 to 29 with the remaining students evenly distributed across the 18 to 23, 3 to 35, and 36 and above categories (23%, 21%, and 21%, respectively). As shown in Table 3.2, 80% of the participants were female and 20% were male. The majority (52%) of participants reported their ethnicity as African American, 36% reported White, and 12% indicated other ethnic groups or failed to report (Table 3.3).

Table 3.1

Age of the Participants

Category	Frequency	Percent	Non-Example Group	Worked Example Group
18-23	14	23%	7	7
24-29	21	34%	9	12
30-35	13	21%	7	6
36 or above	13	21%	7	6

Table 3.2

Gender of the Participants

Response	Frequency	Percent	Non-Example Group	Worked Example Group
Female	48	80%	26	22
Male	12	20%	4	8

Table 3.3

Ethnicity of the Participants

Category	Frequency	Percent
African American	32	52%
Asian / Pacific Islander	0	0%
Hispanic, Latino, or Mexican American	0	0%
White	22	36%
Other	1	2%
N/A	6	10%

Student Experiences with Online Learning

The data presented in Table 3.4 reveals the level of experience the participants have had with online learning and Table 3.5 demonstrates the extent to which it has been integrated with traditional on-campus learning. Of the participants, 90% had prior learning experiences by reporting that they had taken online courses in the past. Two-thirds of the participants reported taking a face-to-face course concurrently with their enrollment in online courses. The remaining third were enrolled exclusively in online courses.

Table 3.4

Previous Online Courses

Response	Frequency	Percent	Non-Example Group	Worked Example Group
Yes	54	90%	28	26
No	6	10%	2	4

Table 3.5

Current Enrollment in On-Campus Courses

Response	Frequency	Percent	Non-Example Group	Worked Example Group
Yes	40	67%	22	18
No	20	33%	8	12

Academic Majors of the Participants

Participants involved in this study covered a wide cross-section of academic majors. Table 3.6 reported the academic majors of the participants. Of the participants, 33% had declared majors of general business, 18% had declared majors in management information systems (the academic area of the course), 11% had declared majors in Accounting, with the remainder distributed across no fewer than seven other majors. Roughly two-thirds of the participants had declared majors in a business-related field.

Table 3.6

Academic Majors of the Participants

Major	Frequency	Percent	Major	Frequency	Percent
General Business	20	33%	Paralegal Studies	2	3%
MIS	11	18%	Management	1	2%
Accounting	7	11%	Psychology	1	2%
Nursing	4	7%	Marketing	1	2%
Other	4	7%	Undeclared	1	2%
Education	3	5%	N/A	6	10%

Debate Procedure

One online debate was offered throughout the course of the semester. The debate began on a Friday morning and concluded the following Thursday evening -- providing a debate period of over six days. The debates were constrained through the use of labels that must be inserted into the subject heading of every message. The labels represented categories (see Table 3.7) similar to those proposed by Jeong (2004).

Table 3.7

Categories of Labels Used in Online Debates

Label	Category	Definition
ARG	Argument	The message supports or opposes a given issue or position.
EVID	Evidence	The message provides evidence, examples, studies, or personal experiences to support a given position.
CRIT	Critique	The message identifies flaws or weaknesses in an opponent's response.
QUES	Question	The message is asking a question for the purpose of clarification.

Also, in accordance with Jeong's (2004) message label categories, team membership was identified by adding an **o** to the end of the label if the message were posted by a member of a team opposing the issue. Conversely, an **s** was added to the end of the message label if posted by a member of a team supporting the issue. For instance, if a student was a member of a team assigned to oppose the statement, "Internet filtering software should be implemented in our schools," he or she might use the following subject heading:

*ARGo Filtering blocks legitimate health information***.**

Then, within the body of the message, he or she would make an argument about the issue of Internet filtering software preventing students from accessing otherwise harmless Web sites about health.

At the onset of the semester, all students were provided basic instructions (Appendix D) about online debates including an overview of the online debate process, descriptions of the message labels, and brief examples of how to use each label. Although no limits were placed on the number of posts made per each debate, a rubric (Appendix E) was provided to ensure the students are aware of the manner in which their grade would be determined.

Worked Debate Example

To conduct the debates, the students in this study were randomly assigned to two groups. To prepare the students for the upcoming debates, the teams in the first group (i.e. control group) were exposed to nothing more than basic instructions and rubric. The teams in the second group were given the opportunity to also see a worked example of an actual debate (Figure 3.1) provided on the discussion board in addition to the basic instructions and rubric.

The worked debate example afforded these students the opportunity to click on links, collapse and expand threads, and interact with the discussion forum just as they will in the real debate to follow. The worked example was physically integrated into the discussion board so as to limit the effects of split attention and the instructions were provided so that the demonstrated debate did not have to be understood in isolation (Sweller et al., 1998). The teams in the control group were given no preparation beyond the instructions. Previous research indicated that this format held the potential to reduce the cognitive load of the students (e.g. Atkinson et al., 2000; Hirsch et al., 2004; Li, 2005; Sweller et al.), assist in message exchange (Jeong, 2004; Jonassen & Remidez, 2005) particularly critical exchanges, allow the students to exercise greater levels of self-control due to the effects of giving the student control over his or her navigation (Clark, 2003; Scheiter & Gerjets, 2007), and improve the social skills exhibited by the participants (Rummel et al., 2009). The provision of examples have also been found to improve student perceptions of instructional activities by making students feel better prepared (e.g. Bozarth et al., 2004; de Bruyn, 2004; Northrup, 2002; Palloff & Pratt, 2003; Rummel et al.) and in greater control of their learning (Reissetter & Boris, 2004). Student perceptions of readiness were cited as necessary for a productive learning environment for the student by Fogerson (2005).

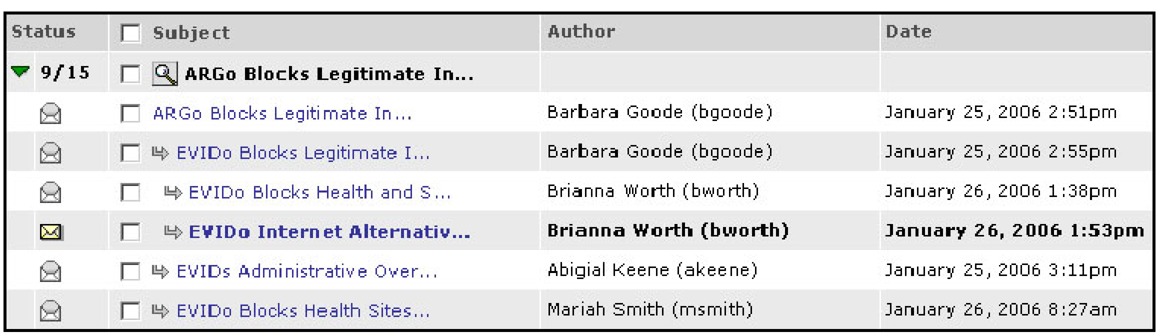

Figure 3.1

Worked Example Screen Shot

The worked example was from a debate conducted by graduate level students at another university who had previously participated in an online debate and had been given feedback from their instructor. These graduate students served as the de facto experts who modeled the expected behaviors within the worked example. To ensure the quality of the worked example and its appropriateness in preparing learners, two levels of review were performed. First, it was selected due to its overall quality by a faculty member with vast instructional and research experiences related to online discussion, online argumentation, and transcript analysis.

Secondly, the data analyses that were to be performed on the online debates were also performed on the worked example. Where possible or appropriate, the results were compared to those studies that informed this research. The sections to follow demonstrate the efficacy of the provided worked example in modeling the behaviors of participation, higher-order cognitive skills, and critical event sequences.

Active and Sustaining Participation

Data were collected about the variables (e.g. post frequency, post length, social cue usage, linguistic qualifier usage, and linguistic intensifier usage) that measured behaviors leading to a more active online debate environment. Eight students participated in the debate demonstrated in the worked example.

Table 3.8 described the participation of the students in the worked example debate as well as made comparisons to Fahy's (2002) research on linguistic qualifier and intensifier usage in computer conferences. There were a total of 11 postings per student. The worked example debate generated over 4100 words, 38 linguistic qualifiers, 21 linguistic intensifiers, and 18 social cues.

Table 3.8

Mean Numbers for Participation Variables in Worked Example and Comparison to Fahy (2002)

	Worked Example	*Fahy (2002)*
Variables	M	M
Posts	11.00	n/a
Words per Post	47.24	n/a
Social Cues	4.33 [a]	n/a
Linguistic qualifiers	9.14 [a]	13.70 [a]
Linguisitc intensifiers	5.29 [a]	5.90 [a]

[a] Number of occurrences per 1000 words used

Cognitive Behaviors

The worked example debate transcript was evaluated by two coders with each posting being placed into one of the five levels of Henri's cognitive framework. Training was performed and the interrater reliability (see page 86) between the coders was found to be very high. In the worked example debate, there were a total of 105 coded messages from the 88 student postings. Three of the messages were coded as not applicable to any of the levels. Table 3.9 described the distribution of messages across the levels of the cognitive framework.

For more meaningful data analysis, Hara et al. (2000) reduced the five levels of Henri's model into two levels: higher-order and lower-order thinking skills. The two clarification levels were related to lower-order thinking skills and the remaining levels were related to higher-order thinking skills. Table 3.10 described the distribution of messages in the worked debate example across the two levels and made comparisons to Hara et al.

Table 3.9

Frequencies of Coded Messages in Worked Example
and Comparison to Hara et al. (2000)

	Worked Example		*Hara et al. (2000)*	
Level	N	%	N	%
Elementary Clarification	21	20.0%	26	14.1%
In-depth Clarification	18	17.1%	20	10.9%
Inferencing	18	17.1%	40	21.7%
Judgment	28	26.7%	64	34.8%
Application of Strategies	17	16.2%	34	18.5%
Not Applicable	3	2.9%		
Total	105	100%	184	100%

Table 3.10

Pooled Frequencies of Coded Messages in Worked Example
and Comparison to Hara et al. (2000)

	Worked Example		*Hara et al. (2000)*	
Variables	N	%	N	%
Lower-order Thinking	39	38.2%	46	25.0%
Higher-order Thinking	63	61.8%	138	75.0%

Event Sequences

The debate message labels were used to identify the function of each message and were loaded into Jeong's (2005a) *Discussion Analysis Tool (DAT)*. DAT used these labels to determine the transitional probability that one type of message would draw a response of another type. Table 3.11 presented the frequencies and transitional probabilities for the worked example debate. The worked example debate generated 73 event sequences.

The results of the generated probabilities indicated that the worked example debate was rich in evidentiary and critical event sequences. Among those event sequences occurring 5 or more times, the most probable event sequence was .50 the ARG -> EVID sequence meaning that postings functioning as EVID postings were made in response to half of all postings functioning as ARG postings. The second and third most probable event sequences were the EVID->CRIT (.45) and CRIT->ARG (.28) sequences.

Table 3.11

Frequency and Transitional Probabilities Matrix

	ARG		CRIT		EVID		QUES	
Label	Freq	Prob	Freq	Prob	Freq	Prob	Freq	Prob
ARG	10	.20	14	.28	25	.50	1	.02
CRIT	0	.00	4	.36	5	.45	2	.18
EVID	1	.00	0	.20	1	.70	0	.10
QUES	0	.50	2	.00	1	.50	7	.00

Student Perception Instrument

A survey instrument (Appendix C) was administered online through the WebCT courseware package. The survey instrument consisted of two parts. Part I sought students' demographic information such as gender, age, major, etc. Part II of the survey contained nineteen declarative statements about the student's perception about their online debate experience. Using Likert scale responses, the students indicated the extent to which they agree or disagree with the statements by clicking on one of the following options: strongly agree, agree, neither agree nor disagree, disagree, or strongly disagree. To minimize the effects of respondents' failure to completely evaluate the questions, several items were reversed from affirmative to negative statements.

Instrument Validity and Reliability

Since the present study is the first to utilize this instrument, validity and reliability data were collected through four semesters of pilot study. Fraenkel and Wallen (2006) defined validity as the "appropriateness, meaningfulness, correctness, and usefulness of the inferences a researcher makes" (p. 150) and called it the most important consideration in preparing an instrument.

The evaluation instrument was designed to measure three primary constructs: *preparedness*, *productive learning environment*, and *learner control* (Table 3.12). The validity of these constructs was established through a foundation in the related literature. Items related to the relative importance of preparing learners to participate in an online course were grounded in research where learners expressed the desire to have clearly stated expectations from instructors (e.g. Carswell et al., 2000; Reisetter & Boris, 2004) and to see examples of good work (e.g. Kawachi, 2002; Sorenson & Baylen, 2004). Also lending validity to this construct were those studies showing that feelings of fright, intimidation, fear, and anxiety were eased given experience with a distance learning exercise (e.g. Bocchi, Eastman, & Swift., 2004; Conrad, 2002; LaPadula, 2003; Perrault, Weldman, Alexander, & Zhao, 2002). Serving as a bridge between the productive learning environment and preparedness constructs is Fogerson's (2005) pragmatic view that satisfaction reflects student perception of the effectiveness and quality of the learning environment and that positive perceptions validate the reasons for offering distance education in the first place.

Table 3.12

Student Perception Domains

Domain	Definition
Preparedness	Statements related to the extent that a student feels prepared to participate in online debate.
Productive Learning Experience	Statements related to the extent that a student feels eager to participate in debate and feels as though it will be a satisfying experience.
Learner Control	Statements related to the extent that a student feels in control of his or her learning within the online debate environment.

The validity of items related to students perceiving the course and online debate to be productive learning environments was demonstrated by studies showing that student satisfaction is "key in producing positive learning outcomes and continuance in a course"

(Conrad, 2002; Menlove & Lignugaris, 2004). The third and final construct being measured was grounded in research showing a relationship between higher levels of learner control and positive learning outcomes and conversely, lower levels of learning control and negative learning outcomes (Clark, 2003; Scheiter & Gerjets, 2007).

The content-related validity was established through a review by a panel of judges capable of offering informed opinions about the adequacy of the instrument (Fraenkel & Wallen, 2006). The expert judges reviewed every question to make certain that every question was content appropriate and clearly understandable. The statement asking the student to indicate their level of agreement or disagreement was removed from every question in order to limit the effects of learner fatigue. Three questions were removed from the survey because they were deemed not to adequately measure the desired content. In addition, some of the demographic questions were reworded.

To ensure the reliability of the instrument, measures of internal consistency were calculated for each of the three types of perceptions. Each type of perception had sufficient internal consistency (Table 3.13) as a result of exceeding the lowest acceptable Cronbach's alpha of .70 recommended by Hair et al. (2006).

Table 3.13

Construct Validity and Reliability of Student Perception Domains

Domains and Items	Cronbach's alpha	Study
Domain 1: Preparedness Clear expectations of course Enough information to patricipate in debate Adequately prepared to participate in online debate Apprehension about course * Confidence about participation in debate Unprepared to participate in online debate * Uncomfortable navigating through online debate * Anxiety about participating in the debate *	0.920	Bocchi et al., 2004; Carswell et al., 2000; Conrad, 2002; Fogerson, 2005; LaPadula, 2003; Perrault, Weldman, Alexander, & Zhao, 2002; Reisetter & Boris, 2004; Kawachi, 2003; Sorenson & Baylen, 2004
Domain 2: Productive Learning Environment Enthusiasm about course Satisfied with the online debate experience Eager to participate in online debate Online debates will hurt learning * Online debates will further learning	0.783	Conrad, 2002; Fogerson, 2005; Menlove & Lignugaris, 2004
Domain 3: Learner control Confused about online debate * Control of learning within learning Confused about course Out of control in debate *	0.719	Clark, 2003; Scheiter & Gerjets, 2007

* item reverse scored

Data Collection

Data were collected over two semesters. Data were collected at various points throughout each semester about the student perception instrument and once throughout each semester about the online debate content. The data were then collectively analyzed in the areas of student perception, content, and event sequence.

Student Perception Instrument

During each semester, the first data were collected through the survey instrument (Appendix B) before the debate and before making the worked examples available to the treatment group. Its purpose was to compare differences between the groups before the exposure to the worked examples began. The second survey (without the demographic questions) was administered after the students in the treatment group reviewed the worked example. The third and final survey (without the demographic questions) instrument was administered again after the debate with the dual purpose of examining if student perceptions changed after experiencing a real debate and monitoring the feeling and perceptions the students took into subsequent online courses.

Content Analysis

In addition to the surveys, data were collected about the debates themselves in order to assess their quality. Each student's postings were compiled into text transcripts. Each transcript were coded based on exposure to the worked example. Example exposure was coded at two levels: 0 (control group) and 1 (having been prepared through the worked example).

Evaluating the quality of discussion presented a unique challenge as no singular, universally agreed-upon system for assessing discussion quality exists. As a result, most researchers have relied on a hybrid approach selecting and modifying evaluative systems from previous research or creating their own. Rourke et al. (2001), for instance, identified nineteen different studies with content analysis schemes -- no two of which utilized the same analysis scheme. In describing the arbitrary nature of coding messages within their own content analysis, Gunawardena, Lowe, and Anderson (1998) explained that "a degree of subjectivity in doing this type of analysis" is unavoidable because researchers are "clearly influenced by their own conceptual frameworks and cultural knowledge" (p. 4).

In this study, a three-pronged approach was taken to operationally assess the online debate quality. Each prong had a strong foundation in the research that informed the current study. In particular, each approach was among those recommended by Hara et al. (2000) and Henri (1992).

Table 3.14

Approaches to Measure Debate Quality

Approach	Variable	Study
Active and Sustaining Participation	Number of postings	Herring (1993); Jeong & Davidson-Shivers (2003)
	Words per post	Barrett & Lally (1999); Herring (2000)
	Linquistic qualifiers	Fahy (2002); Jeong (2005)
	Lingusitic intensifiers	Fahy (2002)
	Social cues	Hara et al. (2000)
Cognitive Skill	Codified levels based on Bloom's Taxomomy	Hara et al. (2000); Kanuka, Rourke, & Laflamme (2007); McKlin, Harmon, Evans, & Jones (2002)
Electronic Interaction Patterns	Event sequences	Fahy, Crawford, & Ally (2001); Jeong (2003); Jeong & Joung (2004); Jeong (2005)

The first approach (i.e. active and sustaining participation) was a measure of the student's level of participation as well as a measure of the words or phrases that typically encourage the participation of others. The second approach (i.e. cognitive level) was to code the cognitive skills displayed in the student messages. The final approach (i.e. electronic interaction patterns) was to measure the patterns of electronic interaction. Table 3.14 showed the resulting variables of each approach along with any previous studies identified to have used a similar analysis.

Active and Sustaining Participation

The first approach in assessing the quality of discussion was to measure behaviors that lead to a more active online debate environment. One set of measures was to simply count the student's debate postings and to calculate the average number of words used per posting. The second set of measures was to count those words or phrases that either promote or limit the participation of others. Finally, the number of times each student used the most common qualifiers or intensifiers (Table 3.15) identified by Fahy (2002) and Jeong (2005b) was counted as well as the number of social cues used. Similar to methods employed by Fahy (2002) and Jeong (2005b), these counts were converted to frequency of usage per 1000 words.

Table 3.15

Variables Measuring Active and Sustaining Participation

Variable	Linguistic Qualifiers	Linguistic Intensifiers	Social Cues
Defintion	A word or group of words that tend to sustain online discussion (Fahy, 2002).	A word or group of words that add emphasis. These words typically limit discussion in an online discussion (Fahy, 2002).	A word or group of words not related to the debate content. These words are often used to acknowledge others (Hara, Bonk, & Angeli, 2000).
Examples	If But I think May/might Though Often Probably	Very Only Always Every Never	Feeling (e.g., "I feel great") Greeting (e.g, "Hello") Closure (e.g., "I'll post more later") Icons (e.g., :-)) Compliments ("Good point.") Using specific name ("John")

Cognitive Level

The second approach in assessing the quality of discussion was to evaluate the cognitive level of the student's messages. According to Hara et al. (2000), exploring the cognitive level of student postings assisted in evaluating the quality because of its relation to reasoning ability, critical thinking skills, and problem solving skills. Going back to the landmark work by Henri (1992), many previous studies created analytical systems based at least, in part, on the cognitive domain of Bloom's taxonomy (e.g. Hara et al.; Kanuka, Rourke, and Laflamme, 2007; McKlin, Harmon, Evans, & Jones, 2002).

These systems tended to place an emphasis on distinguishing the higher levels of cognitive activity in evaluating information from simple recall or statement of fact. Henri also placed a higher value on student ability to evaluate and organize data.

Table 3.16

Analysis Framework: Cognitive Skills

Reasoning Skills	Definition	Indicators
Elementary clarification	Observing or studying a problem, identifying its elements, and observing their linkages in order to come to a basic understanding.	Identifying relevant elements Reformulating the problem Asking a relevant question Identifying previously stated hypotheses Simply describing the subject matter
In-depth clarification	Analyzing and understanding a problem to come to an understanding which sheds light on the values, beliefs, and assumptions which underlie the statement of the problem.	Defining the terms Identifying assumptions Establishing referential criteria Seeking out specialized information Summarizing
Inferencing	Induction and deduction, admitting or proposing an idea on the basis of its link with proportions already admitted as true.	Drawing conclusions Making generalizations Formulating a proposition which proceeds from previous statements
Judgment	Making decisions, statements, appreciations, evaluations and criticisms.	Judging the relevance of statements Making value judgments Judging inferences "I agree, disagree, …."
Application of strategies	Proposing co-ordinated actions for the application of a solution, or following through on a choice or a decision.	Making decisions, statements, appreciations, evaluations, and criticisms.

Note: Hara et al., 2000 (p. 125)

For this study, Hara et al.'s (2000) analytical framework for assessing cognitive thought was used to evaluate the quality of student postings. The model was modified from Henri (1992) and assigned student discussion into one of five levels: *elementary* *clarification*, *in-depth clarification*, *inferencing*, *judgment*, and *application of strategies*. In describing the relationship of Henri's model, Hara et al. indicated that the two clarification levels demonstrated lower-order thinking skills and the remaining three levels demonstrated higher-order thinking skills. In accordance with the Hara et al. model, the paragraph will be the unit of analysis.

Electronic Interaction Patterns

Jeong's (2005a) *Discussion Analysis Tool* (DAT) was used to analyze the sequence of the messages. The message headers of student postings were entered into DAT (Figure 3.2).

	A	B	C	D	E
1	ARG	1	. ARG MediaIsButAMereVehicle		
2	EVID	2	. . EVID MediaIsButAMereVehicle		
3	BUT	3	. . . BUT RelativityTheoryOldToo		
4	BUT	4	. . . . BUT RelativityTheoryOldToo		
5	BUT	2	. . BUT Whataboutemotions?		
6	EVID	2	. . EVID DistEdEffectiveAsFF		
7	BUT	2	. . BUT Mediaamerevehicle		
8	EVID	2	. . EVID MooreConcurs		
9	EXPL	3	. . . EXPL MediaSelectionComesAfterInstructionalStrategy		
10	BUT	3	. . . BUT WellChosenEffective		
11	BUT	4	. . . . BUT SupportingResearch		
12	BUT	2	. . BUT Mediaismorethenamerevehicle		
13	BUT	3	. . . BUT SupportingEvidence?		
14	BUT	2	. . BUT LearningNotSimplyAPassiveResponseToDeliveryMethod		
15	ARG	1	. ARG Standards for teaching		

Figure 3.2

Screen Shot of Jeong's (2005a) Discussion Analysis Tool

DAT was used to first compute the rate that each category of response elicits one or more responses. More specifically, DAT determined the frequency to which each message label elicits responses based on the message label used in the response (Figure 3.3). For example, what was the frequency of students using the EVID label in response to ARG postings (ARG → EVID)?

Figure 3.3

Example of ARG -> EVID Event Sequence

Variables

The independent variable in all data analyses was the group membership of the student. Group membership was coded at two levels: 0 (control group) and 1 (a student having been prepared through the worked example). Students were randomly assigned to these groups.

The dependent variables generated by the survey instrument were the distributions of responses indicating strength of agreement with items related to the domains levels (e.g. preparedness, productive learning environment, learner control) of student perception. A content analysis was used to measure behaviors indicative of active and sustaining participation (e.g. frequency of postings, linguistic qualifiers, linguistic intensifiers) as well as usage of the different cognitive levels identified in Table 3.14. The dependent variable produced by DAT for the event sequence analysis were measurements of the relative frequency of each type of message category exchange.

Data Analysis

A number of different statistical tests (Table 3.17) were conducted to examine differences in the data collected between the two groups.

Table 3.17

Statistical Methods Employed for Data Analysis

Question	Dependent Variable(s)	Method of Analysis
1a.	Mean score of responses to preparedness domain items	Repeated Measures MANOVA
1b.	Mean score of responses to productive learning environment domain items	Repeated Measures MANOVA
1c.	Mean score of responses to learner control domain items	Repeated Measures MANOVA
2a.	Number of posts Length of posts Number of social cues Number of linguistic qualifiers Number of linguistic intensifiers	MANOVA for differences in combined DV and ANOVAs for differences in individual DV
2b.	Distribution of cognitive behaviors across student	Chi-squared test of independence
2c.	Event sequence pairs	Z-scores generated by Jeong's *Discussion Analysis Tool*

For each of the three surveys conducted, descriptive statistics were gathered and a repeated measures MANOVA was performed to test for differences between the groups in their strength of agreement with items related to the *preparedness*, *productive learning environment*, and *learner control* domains.

To analyze the content of the debate transcripts, multiple statistical techniques were used. MANOVA was used to test for differences between the groups on the combined dependent variables of number of posts, length of post, linguistic qualifier usage, linguistic intensifier usage, and social cue usage. ANOVA was then used to test for differences between these variables individually. A chi-squared test of independence was conducted to test for differences in the distribution of higher- and lower-order thinking skills exhibited in posts made by each group of student. Jeong's DAT software program was used to compute the relative frequency in which each event sequence pair occurred. In addition, a z-score for each of the possible event sequence pairing was also calculated based on its relative frequency to determine whether an event pair occurred more or less frequently than expected.

CHAPTER IV

DATA ANALYSIS AND RESULTS

The purpose of this research was to explore how students' perceptions of online debate and the content of their debate postings differed depending on exposure to a worked example before participating. The study utilized both descriptive and inferential statistics in order to explore the population and determine those variables in which statistically significant differences existed between the control and treatment groups. This chapter presents the results of this study and is divided into five sections: (a) statistical analysis of the population, (b) analysis of student responses to items on the perception survey, (c) analysis of behaviors indicating a student's active participation and sustaining the participation of other students, (d) analysis of cognitive behaviors, and an (e) event sequence analysis of the postings.

Statistical Analysis of the Population

Two levels of statistical analysis were performed to ensure similarities between the control and treatment group. Starting with demographic characteristics, the control group and the group being exposed to the worked example were similar with respect to gender, $\chi 2$ (1, N = 61) = 1.667, p = .197, age of participants, $\chi 2$ (3, N = 61) = .566, p = .904, racial/ethnic composition, $\chi 2$ (3, N = 61) = 2.795, p = .424, and previous experience with online courses, $\chi 2$ (1, N = 61) = .669, p = .414.

The second level of analysis sought to ensure that there were no statistically significant differences between the groups on their self-reported perceptions on the three domains being evaluated. The student perceptions were collected from the survey administered to the students before the treatment group was exposed to the worked example. Independent sample t-tests revealed no significant differences between the groups on their mean scores for survey items related to the preparedness domain, $F(1, 59) = .659$, $p = .420$, the productive learning environment domain, $F(1,59) = 1.248$, $p = .260$, or the learner control domain, $F(1, 59) = .251$, $p = .610$.

Analysis of First Research Question

Did the worked example have a significant effect on student's self-reported perceptions about the online debate? To answer the first research question, descriptive and inferential statistics about items related to the *preparedness*, *productive learning environment*, and *learner control* domains were reported for each of three surveys. Survey one (i.e. pre-treatment survey) was administered prior to exposing the treatment group to the worked example so that all participants were equivalent in terms of their exposure to the worked example. Survey two (i.e. post-treatment survey) was administered after the treatment group was exposed to the worked example. Survey three (i.e. post-debate survey) was administered after the completion of the online debate.

The sections to follow answer the first research question regarding differences in student perception between the treatment and control group as well as across surveys.

Statistical Analysis of the Student Perception Survey

A repeated measures MANOVA was performed to compare the mean scores from pre-treatment, post-treatment, and post-debate surveys on the *preparedness*, *productive learning environment*, and *learner control* domains. The results of the repeated measures MANOVA demonstrated that the combined domain means within the entire sample changed across surveys, Wilks' $\Lambda = .856$, $F(6, 39) = 2.319$, $p = .035$, $\eta^2 = .30$, but that exposure to the worked example did not significantly affect the combined domain mean across surveys, Wilks' $\Lambda = .998$, $F(3, 42) = .022$, $p = .996$, $\eta^2 = .002$. Subsequent univariate tests (Table 4.1) revealed only a significant change in the productive learning environment domain across surveys ($F(2, 44) = 4.298$, $p = .002$, $\eta^2 = .089$), but no significant changes in perception were found on the preparedness ($F(2, 44) = 1.464$, $p = .237$, $\eta^2 = .032$) and learner control domains ($F(2, 44) = 1.448$, $p = .241$, $\eta^2 = .032$).

To determine the surveys on which the mean score of the productive learning environment domain were significantly different, a modification of Tukey's HSD test designed for repeated measures analysis was performed (Gravetter & Wallnau, 2004). The results revealed that the mean score of the productive learning environment differed significantly for the entire sample between the pre-treatment survey and the post-debate survey, but not between either the pre-treatment and post-treatment survey or the post-treatment and post-debate survey.

The survey and group membership, however, did not interact significantly on the combined domain mean, Wilks' $\Lambda = .983$, $F(6, 39) = .241$, $p = .962$, $\eta^2 = .041$. Univariate tests (Table 4.2) revealed no significant changes in the preparedness ($F(2, 44) = .210$, $p = .811$, $\eta^2 = .005$), productive learning environment ($F(2, 44) = .616$, $p = .542$,

$\eta^2 = .014$), and learner control ($F(2, 44) = .229$, $p = .796$, $\eta^2 = .005$) domains across surveys with group membership being used a factor.

Table 4.1

Changes in Domain Mean Scores From Pre-Treatment Survey
to Post-Debate Survey

Domain	*Survey*	M	SD	F(2,44)	Sig.
Preparedness	Pre-Treat	4.08	0.71	1.464	.237
	Post-Treat	4.07	0.58		
	Post-Debate	3.97	0.67		
Prod Environment	Pre-Treat	4.34	0.56	4.298	0.02 *
	Post-Treat	4.20	0.76		
	Post-Debate	4.12	0.81		
Learner Control	Pre-Treat	4.06	0.66	1.448	.241
	Post-Treat	4.19	0.75		
	Post-Debate	4.00	0.83		

* Significant at the 0.05 level

Table 4.2

Changes in Domain Mean Scores From Pre-Treatment Survey
to Post-Debate Survey by Group Membership

		Non-Example		*Worked Example*			
Domain	*Survey*	M	SD	M	SD	F(2,44)	*Sig.*
Preparedness	Pre-Treat	4.05	0.71	4.13	0.58	0.210	.811
	Post-Treat	4.09	0.57	4.07	0.59		
	Post-Debate	3.94	0.64	3.99	0.73		
Prod Environment	Pre-Treat	4.32	0.55	4.37	0.58	0.616	.542
	Post-Treat	4.23	0.87	4.17	0.64		
	Post-Debate	4.07	0.96	4.18	0.63		
Learner Control	Pre-Treat	4.02	0.65	4.10	0.69	0.229	.796
	Post-Treat	4.21	0.73	4.17	0.78		
	Post-Debate	3.96	0.81	4.06	0.88		

Summary of First Research Question

A repeated measures MANOVA was performed to compare the mean scores from pre-treatment, post-treatment, and post-debate surveys on the *preparedness*, *productive learning environment*, and *learner control* domains. The results revealed that the time of survey significantly affected the combined dependent variable of domain mean scores, but that exposure to the worked example did not interact significantly on the combined dependent variable or on any of the individual domain means. Univariate tests demonstrated that the mean score for the productive learning environment domain differed significantly within the entire sample across surveys, but not between members of the control group and the group exposed to the worked example. The failure to report significant differences on the productive learning environment mean score between groups was due to a similar downward trend in self-reported scores from the pre-treatment survey to the post-treatment survey between members of the control group and those exposed to the worked example (Figure 4.1).

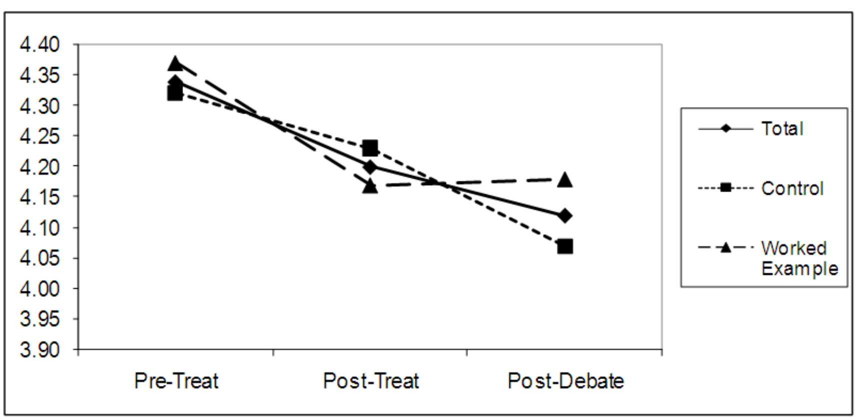

Figure 4.1

Trend of Productive Learning Environment Domain
Mean Score by Group

Analysis of Second Research Question

Did the worked example have a significant effect on the students' learning behaviors in the online debate? To answer the second research question, descriptive and inferential statistics on were reported about the three major elements of student content: (a) active and sustaining participation, (b) cognitive skills, and (c) event sequences. The sections to follow answer the following research questions regarding student behaviors indicative of student participation and sustaining the participation of others.

Active and Sustaining Participation

Were there significant mean differences in the number of behaviors indicative of active participation and sustaining participation (number of postings, words per post, social cue usage, linguistic qualifier usage, and linguistic intensifier usage) for students being exposed to the worked example and those students who were not? The sections to follow answer the research question regarding the behaviors indicative of active participation and sustaining the participation of others within student postings.

Descriptive Statistics about Student Participation

Sixty-four students participated in the 12 debates. Table 4.3 described the participation of the students in these debates. There were a total of 271 postings leading to an average of 4.16 postings per student per debate. The postings generated over 28,000 words, 299 linguistic qualifiers, 96 linguistic intensifiers, and 46 social cues.

Table 4.3

Mean Numbers for Variables as a Function of Group Membership

Variables	*Non-Example*		*Worked Example*	
	M	SD	M	SD
Posts	3.45	3.00	5.07	2.28
Words per Post	115.40	81.02	94.45	40.60
Social Cues	1.67 [a]	2.58	2.59 [a]	3.81
Linguistic qualifiers	8.19 [a]	6.67	11.20 [a]	7.90
Linguisitc intensifiers	2.90 [a]	3.35	5.32 [a]	4.38

[a] Number of occurrences per 1000 words used

Statistical Analysis of Student Participation

A one-way MANOVA was carried out in order to examine the effects of the worked example exposure on the combined dependent variable of post frequency, post length, social cue usage, and the usage of linguistic qualifiers and intensifiers. Univariate ANOVA tests were run to explore individual mean differences between the control group and the group being prepared through exposure to a worked debate example.

Three cases (one from the worked example group and two from the control group) were identified to be multivariate outliers through the calculation of Mahalanobis distances and were eliminated from the analysis. The assumption of univariate normality was violated for four of the five variables. Attempts at transforming the variables did not substantively improve univariate normality. Graphical checks of normality revealed these violations to be only small deviations from normal. For all variables, the non-normality stemmed almost entirely from positive skewness and not from the presence of outliers. Given these facts and the size of the sample, MANOVA proved robust to these violations (Hair et al., 2006).

Examinations of bivariate scatterplots revealed linear relationships among all combinations of dependent variables meaning no further transformations were necessary. A Box's Test of Equality Covariance Matrices ensured that the assumption of homoscedasticity was met. A Levene's Test of Equality of Error Variances proved univariate homogeneity for four of the five variables. According to Mertler and Vannatta (2005), violations of this assumption were not fatal. Likewise, Hair et al. (2006) stated violation of these assumptions were not fatal given the similarity of sample sizes. A correlation matrix (Table 4.4) was generated to ensure that multicollinearity was not a threat to the analysis. No correlations among variables were high enough to be problematic.

Table 4.4

Correlation Matrix for Variables

	1	2	3	4
1. Posts				
2. Words per Post	-.166			
3. Social Cues	.144	-.172		
4. Linguistic qualifiers	.139	-.029	.256 *	
5. Linguisitc intensifiers	.133	-.101	.112	.050

** Significant at the 0.05 level

A MANOVA was employed to analyze the effects of the exposure to the example on student behaviors related to the active and sustaining participation variables. The combined dependent variables of number of posts, words used per post, social cue usage, linguistic qualifier usage, and linguistic intensifier usage were significantly affected by exposure to the example (Wilks' $\Lambda = .809$, $F(1, 28) = 2.597$, $p = .035$, $\eta^2 = .206$).

To examine the impact of the use of examples on the dependent variables individually, follow-up ANOVA tests were performed. The results of the univariate ANOVA tests reveled that example exposure had a significant effect on the usage of number of postings ($F(1, 59) = 5.633$, $p = .021$, $\eta^2 = .062$) and linguistic intensifiers ($F(1, 59) = 5.863$, $p = .019$, $\eta^2 = .096$). Students being exposed to the worked example (M = 5.07, SD = 2.28) made a greater number of posts than those students in the control group (M = 3.45, SD = 3.00). In addition, students being exposed to the worked example (M = 5.32, SD = 4.38) made more frequent usage of the linguistic intensifiers than those students in the control group (M = 2.90, SD = 3.35). Its effects on length of postings ($F(1, 59) = 1.611$, $p = .209$, $\eta^2 = .043$), social cue usage ($F(1, 59) = 1.266$, $p = .265$, $\eta^2 = .02$), and linguistic qualifier usage ($F(1, 59) = 2.585$, $p = .113$, $\eta^2 = .045$), however, were found to be non-significant.

Cognitive Behaviors

Were there differences in the distribution of cognitive skills exhibited in the postings of students being exposed to the worked example and the postings of students who were not? The sections to follow answer the research question regarding the cognitive skills exhibited in the student postings.

Interrater Reliability

The debate transcripts were coded by two coders with each posting being placed into one of the five levels of Henri's cognitive framework. The measurable constructs from the framework were discussed with the coders and a pilot subsample was practiced with the coder. To establish the reliability of the ensuing coding, Krippendorff's alpha was calculated. It was chosen due to its ability to take into account chance agreement between coders (de Wever et al., 2006; Neuendorf, 2002; Rourke et al., 2001) and because it is a conservative choice compared to more liberal alternatives such as percent agreement and Cohen's kappa (de Wever et al.). The reliability of the codings were considered good for the control group ($\alpha = .68$), the worked example group ($\alpha = .76$), and for the overall sample ($\alpha = .73$). In those instances in which a message received different codings, a consensus was reached through discussion among the coders.

Descriptive Statistics about Cognitive Behaviors

In the debates, there were a total of 297 messages coded from the 271 student postings. Table 4.5 described the distribution of postings across the five levels of Henri's cognitive framework. Twenty-four of the messages were coded as not applicable. The not applicable messages fell into one of three categories. The first category were "double posts" in which the student submitted the same post twice. The second category were "corrective posts" in which the student made a follow-up post to his or her initial post correcting some element (most often a broken hyperlink) within the initial post. The third category were "off-task posts" in which the student's post had nothing to do with the debate topic. An example would be when one student posted how a cell-phone technology improved society and another student responded with a question about how to acquire that technology.

Table 4.5

Frequencies of Coded Messages by Group

Level	*Non-Example* N	*Non-Example* %	*Worked Example* N	*Worked Example* %	*Total* N	*Total* %
Elementary Clarification	30	20.1%	33	22.3%	63	21.3%
In-depth Clarification	44	29.5%	26	17.6%	70	23.6%
Inferencing	21	14.1%	33	22.3%	54	18.2%
Judgment	26	17.4%	36	24.3%	62	20.9%
Application of Strategies	10	6.7%	13	8.8%	23	7.8%
Not Applicable	18	11.4%	6	4.1%	24	8.1%
Total	149	100%	147	100%	296	100.0%

Table 4.6

Pooled Frequencies of Coded Messages by Group

Variables	*Non-Example* N	*Non-Example* %	*Worked Example* N	*Worked Example* %	*Total* N	*Total* %
Lower-order Thinking	74	56.5%	59	41.8%	133	48.9%
Higher-order Thinking	57	43.5%	82	58.2%	139	51.1%
Total	131	100%	141	100%	272	100.0%

Statistical Analysis of Cognitive Behaviors

Hara et al. (2000) related the two clarification levels of Henri's cognitive framework to lower-order thinking skills and the remaining levels to higher-order thinking skills. Before performing statistical analysis, the distribution of coded messages across the five levels in Henri's framework were reduced to the two identified by Hara et al.: higher-order thinking skills and lower-order thinking skills. Table 4.6 described the distribution of messages across the two levels.

A chi-square test of independence was performed to examine the relationship of exposure to the worked example and the distribution of messages coded as exhibiting either higher-order or lower-order thinking skills. The results of the chi-square test of independence test (Table 4.7) indicated that the distribution of messages coded across the higher-order and lower-order thinking skills levels differed significantly between the worked example group and the control group ($\chi 2$ (1, N = 272) = 5.828, p = .016). The proportion of coded messages indicating higher-order thinking was greater for those students prepared through worked example (58%) than for those students in the control group (44%). Conversely, the proportion of messages indicating lower-order thinking skills was smaller for those students in the worked example group (42%) than for those in the control group (56%).

Table 4.7

Chi-Square: Cognitive Skills by Group

Group	Lower Order	Higher Order	*χ2*	*Sig.*
Non-example	74	59	5.828	0.016 *
(Residual)	(1.2	(-1.2)		
Worked Example	57	82		
(Residual)	(-1.2)	(1.2)		

* Significant at the .05 level

Event Sequence Analysis

Were there differences in the frequencies of event sequences for students being exposed to the worked example and the postings of students who were not? The sections to follow answer the following research question regarding the event sequences of labels used in the headings of student postings.

<u>Descriptive Statistics about Event Sequences</u>

The debate message labels were used to identify the function of each message. Jeong (2003) identified this method of using message labels as accurately predicting the event that occurs in the body of the message. Jeong (2005a) also developed the computerized *Discussion Analysis Tool (DAT)* which was used to accept these labels as inputs and outputting the transitional probabilities between the events. In other words, how often did a message of one type draw a response of another type (i.e. CRIT posting in response to an ARG posting)? These probabilities can be used to relate the structure of the debate to critical thinking. DAT was also used to determine which transitional probabilities were statistically significant.

Table 4.8 presented the frequency matrix generated by DAT for debates conducted by the control group and the group being exposed to the debate examples. Table 4.9 presented the transitional probabilities matrix generated by DAT. The control group debates generated 58 event sequences and the debates conducted by the group exposed to the debate example generated 52 event sequences. The most probable event sequence for the control group debates was .47 for the EVID->EVID event sequence meaning that 47% of all responses to EVID postings in these debates were EVID responses. The second and third most probable were ARG->ARG (.40), and CRIT->CRIT (.40) sequences. The three most probable event sequences for the group being exposed to the debate sample were the CRIT->CRIT (.64), ARG->CRIT (.50), and EVID->ARG (.45) sequences.

Table 4.8

Frequency Matrix

	ARG		CRIT		EVID		QUES	
Label	Non [a]	Ex [b]	Non	Ex	Non	Ex	Non	Ex
ARG	7	4	11	15	10	11	0	0
CRIT	4	0	4	7	2	1	0	0
EVID	4	3	5	4	9	2	0	0
QUES	0	5	0	0	0	0	0	0

[a] Control group not receiving debate example; [b] Group exposed to debate example

Table 4.9

Transitional Probabilities Matrix

	ARG		CRIT		EVID		QUES	
Label	Non [a]	Ex [b]	Non	Ex	Non	Ex	Non	Ex
ARG	.25	.13	.39	.50	.36	.37	.00	.00
CRIT	.40	.27	.40	.64	.20	.09	.00	.00
EVID	.21	.45	.26	.36	.47	.18	.00	.00
QUES	.00	.00	.00	.00	.00	.00	.00	.00

[a] Control group not receiving debate example; [b] Group exposed to debate example

Statistical Analysis of Event Sequences

DAT also produced Z-scores (Table 4.10) for these transitional probabilities that reflected "whether the observed probabilities were higher or lower than the expected probabilities based on a random distribution of responses across each message category" (Jeong, 2006, p. 204).

Table 4.10

Z-Scores Matrix

	ARG		CRIT		EVID		QUES	
Label	Non [a]	Ex [b]	Non	Ex	Non	Ex	Non	Ex
ARG	-.14	**-1.95**	.74	.00	-.34	**1.85**	-.97	-.01
CRIT	1.12	.37	.40	1.02	-1.28	-1.50	-.46	-.01
EVID	-.58	**1.98**	-.91	-1.02	1.03	-.74	1.45	-.01
QUES	-.60	-.01	-.73	-.01	1.29	-.01	-.13	.00

[a] Control group not receiving debate example; [b] Group exposed to debate example;
Significant at .05

These z-scores were used to answer the research question regarding differences in the frequencies of event sequences for students in the treatment and control group. For the control group, the z-scores revealed no transitional probabilities that occurred significantly higher or lower than expected. For the group exposed to the worked example, the z-scores revealed three interactions that occurred significantly higher or lower than expected: ARG->ARG, ARG->EVID, and EVID->ARG. The probability that an ARG posting was followed by another ARG posting was .13 and significantly lower than expected (z-score = -1.95, p = .03). The other significant interactions occurred more frequently than expected. The probability that an ARG posting was followed by an EVID posting was .37 (z-score = 1.85, p = .03) and the probability that an EVID posting was followed by an ARG posting was .45 (z-score = .1.98, p = .02). No other interactions were found to be significant.

Summary of Second Research Question

Each student's debates were compiled and analyzed. Five variables were collected that reflected the student's participation and his or her ability to sustain the participation of others. A MANOVA test was conducted to analyze the effect of example exposure on the combined dependent variable of number of posts, words used per post, social cue usage, linguistic qualifier usage, and linguistic intensifier usage. The results of the test showed that worked example exposure had a significant effect on the combined dependent variable. Specifically, univariate ANOVA tests revealed a significant mean difference in linguistic intensifier usage and number of postings between those in the control group and those being prepared through exposure to the worked example.

Postings from all debates were compiled, analyzed, and coded into one of the five cognitive levels identified by Henri's framework. Descriptive information about the distributions of coded messages was reported by group. Pursuant to the idea presented by Hara et al. (2000), the original five levels were broken down into the broader levels of higher-order and lower-level thinking skills. A chi-square test of independence revealed a significant difference between the groups for the distribution of coded messages across the higher-order and lower-order thinking skill levels with the students being exposed to the worked example making a higher proportion of messages exhibiting higher-order thinking skills. The message labels from each posting in the debates were loaded into Jeong's *Discussion Analysis Tool* (2005a). DAT produced matrices reflecting the relative frequencies of all possible event sequences and the transitional possibility of all possible event sequences. DAT also performed significance testing by calculating z-scores that indicated whether the observed frequency of an event sequence occurred more of less frequently than expected. Based on the results of the z-tests, no interactions were found to be significant for the control group debates. For the debates conducted by the group being exposed to the debate example, two interactions (ARG->EVID, EVID->ARG) were found to occur significantly more often than expected and one (ARG->ARG) was found to occur significantly less often than expected.

CHAPTER V

SUMMARY AND CONCLUSIONS

This chapter offers a summary of the study of the effect of preparing students for an online debate through a worked example and consists of six major areas: (a) an overview of the procedures employed and the principal findings, (b) a discussion of these key findings, (c) implications for utilizing worked examples to prepare students for new or unfamiliar instructional designs, (d) limitations of the study, (e) recommendations for future research, and (f) concluding comments.

Summary of Study

This study examined the effects of preparing students through exposure to a worked example on both student perceptions and the learning behaviors exhibited in online debate. This section summarizes the research methodology employed to test the research questions and concludes with a report of the principal findings.

Procedures

In this study, students were randomly assigned to teams to participate in an online debate exercise conducted on the course discussion board. Half of the assigned students were given basic instructions about how to participate in the debate while the other half were also shown a worked example of an online debate in addition to these basic instructions.

Data were collected from two sources. First, a student perception instrument measuring the areas of *preparedness*, *productive learning environment*, and *learner control* was administered at various points throughout the semester. Secondly, data were collected about the debates and a content analysis was performed. Since no universally-agreed upon system for assessing or operationally defining the quality of online discourse, a hybrid approach consisting of "best practice" methods from existing research was used. Three approaches were used. The first approach was to count the instances of behaviors found to be indicative of a student actively participating in a discussion and those behaviors linked to encouraging or sustaining the participating of others. The second approach was to code the cognitive behaviors exhibited in student postings into one of five levels. The third approach was to examine the electronic interaction patterns as exhibited by the message labels.

Principal Findings

This study generated the following key findings based on responses to the student perception instrument and the content collected from the online debate transcripts.

1. Combined mean scores from the *preparedness*, *productive learning environment*, and *learner control* domains differed significantly within the entire sample across surveys. Exposure to the worked example, however, did not interact significantly with the time of survey to affect mean scores on any of the three domains.
2. Students in the control group and the worked example group differed significantly in those behaviors collectively shown to be indicative of active

and sustaining participation. The worked example group made significantly more posts and used significantly more linguistic intensifiers.

3. Students in the worked example group posted a higher proportion of messages indicating higher-order thinking skills than those students in the control group.
4. There were no statistically significant event sequences measured for the control group. Conversely, DAT found three event sequences for the worked example group that occurred more or less frequently than expected.
    a. Students in the worked example group used the ARG -> ARG event sequence less frequently than expected.
    b. Students in the worked example group used the ARG -> EVID and EVID -> ARG event sequences with greater frequency than expected.

Discussion

This study adds to the existing literature that identified worked examples as a means to greater levels of participation, more critical thinking, and more cognitively-rich higher-order thinking. The sections to follow describe the study's major findings in the context of the literature that informed it.

Student Perception

One purpose of this study was to examine the impact of a worked example on the following domains or areas of student perception: (a) how prepared students felt to engage in online debate, (b) how productive the online debates would be in furthering their learning, and (c) the extent to which students felt in control of their learning. Repeated measures MANOVA revealed a significant main effect in that the collective mean scores of the three domains differed significantly across the three surveys. Generally, the mean for each domain regressed downward from the pre-treatment survey to the post-debate survey.

Examining the results for the *survey by group* interaction indicated no significant effects on the domain means collectively or individually. This finding was key in answering the first research question and its associated sub-questions. The results showed that students who were exposed to the worked example did not differ significantly from the control group in their responses to student perception items. This finding stands in contrast to the research that informed the current study that suggested a link between positive learner perceptions with clarifying learner expectations for complex instructional tasks and providing them with worked examples of these tasks. The prior research generally fell into one of three types.

One type were those studies that surveyed students to determine empirically if students would like to see examples of quality work at the onset of an activity (e.g. Bozarth et al., 2004; Northrup, 2002; Reisetter & Boris, 2004). Carswell et al. (2000) similarly used a survey, but also supplemented their research through interviewing subjects. The second type were those studies that simply theorized that examples would lead to more positive student perceptions (e.g. Palloff & Pratt, 2003; Sorenson & Baylen, 2004). The third type drew a causal comparison between positive student perceptions and student achievement and retention (e.g. Fogerson, 2005; Song, 2003).

The possible reason that exposing students to a worked example before participating in the online debate did not have the discernible positive effect on student perceptions as suggested by the research were methodological differences between the previous efforts and the current study. The present research expanded upon previous research by gauging perceptions of students exposed to a specific type of preparation; in this case the preparation being a worked example. Rummel et al. (2009) utilized a similar procedure and also found no statistically significant difference between the group being exposed to what they called a model condition and the control group receiving no additional pre-instruction or preparation.

There were many instances of the prior research in this area, however, that simply asked the participants to indicate that they would perform better if they were given better preparation, a clearer idea of instructor expectations, and access to worked examples. Other examples of this research provided no empirical foundation, but instead argued the theoretical merits of preparing student through worked example would lead to more positive student perceptions of interactive activities. These studies may lend credence to Gosrky and Caspi's (2005) contention that some widely accepted theories in student interaction have been accepted simply due to their high face validity or Kreijns et al.'s (2002) assertion that interaction has too often been considered a solution without a sufficient grounding in the research.

Among the studies informing this research that related student perception with student readiness, the current study and the research done by Rummel et al. (2009) were the only ones identified that failed to draw a strong relationship between positive student perceptions and their sense learner of preparedness. Unlike these other research efforts into this area, these studies provided actual worked examples and tested to see if observing these examples made students feel more prepared and more positive about the impending activity.

Active and Sustaining Participation

This study explored the impact of worked examples on behaviors that exhibited active participation and sustained the participation of others. An evaluation of the descriptive statistics showed that on four of the five measures, students in the worked example group more closely emulated the modeled behaviors than did students in the control group.

A MANOVA was carried out in order to examine the effects of the worked example exposure on posting frequency, length of post, linguistic qualifier usage, linguistic intensifier usage, and social cues. The results indicated an overall difference among the groups. Testing for group differences on the variables individually showed that students being exposed to worked example made significantly more posts and used significantly more linguistic intensifiers.

Mostly, these findings were expected and supported previous research that indicated similar impacts of worked examples. One possible explanation is the research that finds learners do emulate behaviors modeled through worked examples (e.g. Atkinson et al., 2000; Perdersen & Liu, 2002; Rummel et al., 2009). Beyond modeling, the worked example possibly reduced extraneous cognitive load thus leading to greater participation (e.g. Atkinison et al.; Hirsch et al., 2003; Sweller, 1998, van Gog et al., 2004). Exposure to the worked example offered students the opportunity to move more of the load in learning the debate procedures into long-term memory, thus freeing up working or short-term memory to actually make posts and interact with their colleagues. The worked example also assisted students in navigating the complex hypertext format which limited the amount of time they spent in process-oriented tasks such as overcoming navigational disorientation, searching through the structure, and accessing information (e.g Chen, 2003; Scheiter & Gerjets, 2007; Pedersen & Liu). The end result being students exposed to the worked example were able to spend a greater amount of their cognitive effort on posting content and less time and effort on learning the process.

One of the findings deserves further explanation when viewed in context of the prior research. This finding concerns linguistic intensifiers which have been thought to limit dialogue (Fahy, 2002). Despite posting more frequently, the students begin exposed to the worked example actually made significantly greater use of linguistic intensifiers than did students in the control group. This finding may be an indication of the effect of modeling being greater than the effect of intensifiers to limit dialogue. The number of linguistic intensifiers used by the worked example group (M = 5.32) were almost identical to number of intensifiers (M = 5.29) modeled in the worked example.

Cognitive Skills

This study also examined the impact of a worked example on the distribution of higher-order thinking skills exhibited in student messages. An evaluation of the descriptive statistics revealed that the cognitive behaviors exhibited by the students (58% of messages demonstrated higher-order thinking skills) observing the worked example prior to participating in the online debate more closely resembled the cognitive behaviors modeled in the provided example (62%) than those cognitive behaviors exhibited by the control group (44%). A chi-square test of independence was performed and confirmed that a higher percentage of the messages made by students in the treatment group demonstrated higher-order thinking skills. This significant difference in the proportion of higher-order thinking skills between groups was expected, which supports previous research into worked examples and the various functions they perform in assisting students with complex instructional tasks and new instructional environments.

One such function was the ability of worked examples to reduce the extraneous cognitive load placed on the learner by having to learn the format of the instruction (Sweller et al., 1998). Van Gog et al. (2004) suggested that regardless of the problem, learners inexperienced with an instructional format would utilize weak techniques to solve problems. In a study by Li (2005), worked examples were found to assist in reducing mental effort scores which, in turn, led to improved cognitive measure scores. The online debate utilized in the current study also presented a challenging instructional format requiring some degree of cognitive load to learn. Those students given an opportunity to explore the format through the worked example showed a greater number of higher-order thinking skills than those who had to exercise the mental effort to learn the format of their own accord.

Another explanation arises from the role worked examples play in automating schema construction. Those students given the opportunity to observe the worked example were very much like the chess players that Sweller et al. (1998) wrote about. Experienced chess players organize and categorize numerous different board positions prior to participating in chess matches. The treatment group students in the current study similarly used the worked example to organize different elements (e.g. label construction, message composition, general navigation) of the online debate prior to participation. As a result of automating the construction of the online debate schema, these students were able to move the procedural elements of the debate to long-term memory and utilize a greater amount of their working memory to proposing ideas, making evaluations and criticisms of information, and solving problems. In addition, the manner in which the worked example was presented limited the effects of split attention and the extraneous cognitive load it causes.

By offering students access to all the information contained within the worked example from a single location, students in the treatment group did not have to divide their attention between two locations or channels of information (Sweller et al., 1998). Instructions were also provided so that the worked example did not have to be understood in isolation (Sweller et al.).

Event Sequences

The final approach of this study was to evaluate the influence of exposure to a worked example on the cognitively rich exchanges of critical thought and information. Jeong's (2005a) *Discussion Analysis Tool* was used to evaluate which sequences of message labels were used more or less often than expected by chance. The results indicated that students in the control group used no event sequences more or less often than expected.

There were, however, findings of event sequences occurring significantly more and less frequently than expected for students who were exposed to the worked example prior to participating in the online debate. Occurring less frequently than expected were the ARG->ARG sequence in which a posting featuring an ARG message label were used in response to another posting featuring another ARG message label. Among the different types of event sequences, the ARG->ARG sequence may be the least productive in generating critical thinking, introduction of new evidence, and promoting interaction among discussants. Jeong and Joung (2004) cited students rarely responding "to arguments with evaluation of the argument's accuracy, validity, and relevancy" as problematic due to the fact that the ARG->ARG event sequence occurred more frequently than expected by chance in their study (p. 37).

The results of those findings of event sequences occurring more frequently than expected by chance in the treatment group online debates were more mixed. The finding that both the ARG->EVID and EVID->ARG sequences occurred more frequently than expected by chance confirmed some degree of influence of the worked example exposure on the posting behaviors of the students. Jeong and Joung (2004) used a high number of evidentiary postings as one of their theoretical assumptions for good argumentation to ensue. Oh and Jonassen (2007) found that scaffolding argumentation aided novice learners in generating evidence. The worked example provided to the treatment group in the present study featured a large number of EVID postings. In accordance with those theories that suggested students learners would emulate behaviors seen in a worked example, it appeared that the treatment group in this respect did emulate the behaviors of the experts in the worked example.

The provided worked example also featured a large number of CRIT postings and two of the three most frequently occurring event sequences were the EVID->CRIT and the CRIT->ARG sequences. Jeong (2003) called exchanges of arguments and criticisms triggering mechanisms that help students better understand arguments.   He considered this type of exchange so important that he suggested that instructors utilizing online debate make clear expectations and even promote these argumentative exchanges. A worked example rich in critical exchanges appeared to be a possible solution. Despite the fact that students mimicked the behaviors demonstrated in the worked example in so many other ways (e.g. participation levels and types, cognitive behaviors, use of evidence), they did not use any one of the CRIT-based event sequences significantly more than expected by chance. The worked example in the current study may not have had the level of critical exchange necessary to overcome Jeong's (2003) assertion that many students avoided using the CRIT label due to it appearing to be overly confrontational.

Implications

Although this study provides only a starting point to explore the potential of worked examples to prepare students to participate in an instructionally complex interactive environments, its findings have important educational implications.

The results of this study supported previous research that indicated scaffolds could be used to produce more components of argumentation and learners prepared through expertly modeled worked examples would emulate the behaviors being modeled. The present study has extended the research that informed it in many ways. Some studies have in a general sense professed CSCL environments to suffer from lower participation, lower quality learning outcomes, and lower learning satisfaction (e.g. Joung & Keller, 2004; Kreijns et al., 2002). Many of the reviewed studies have attempted to prove the benefits of clarifying expectations and providing examples (especially worked examples) to students prior to their participation in unfamiliar instructional formats. Excluding Rummel et al. (2009), the research that informed the current study has not related the influence of specific techniques modeled through worked example on key measures, nor has it examined the impact of worked examples on preparing learners for different types of activities. As a result, this gap in the research might leave researchers, course designers, and instructors asking three key questions: Are different types of worked examples more effective in preparing learners than others? Are certain types of behavior more easily modeled than others? Do worked examples hold greater promise for certain types of interactive instructional activities than others?

In addressing these concerns, this study used key measures defined by previous research to operationally assess student perception and the quality of the dialog in an online debate. Its findings established that at least in the context of the online debate format, students exposed to worked examples modeled many of the desired participative and cognitive behaviors, posted more frequently, and posted more information both in support of arguments or to refute claims made by others. The practical recommendation derived from these findings is that instructors and course designers should provide worked examples to students before their participation in online debates and similarly structured constraint-based argumentation activities. Although the implications are only generalizable to other environments using a similarly constructed worked example and constraint-based activity, this study's findings should not be so narrowly construed as to limit future research into other types of worked examples and activities. On the contrary, the present study has laid the foundation and set forth the theoretical assumptions necessary to justify further research utilizing different types of worked example and across different types of activities. If worked examples can be consistently shown to better prepare learning and improve performance, it stands to reason that new and innovative instructional ideas can be more aggressively pursued.

Recommendations for Future Research

The present study utilized participants who were students from a small, liberal arts university in the Southeastern United States. A replication of this study should be performed at other types of institutions that might be reflective of different student populations or settings. Beyond a reflection of different institutional values, future research is also needed to get a more comprehensive and precise understanding of the relationship between worked examples and student perception, preparedness, and performance. Following are recommendations for future research relative to gaining a greater understanding of the nature of worked examples.

This study found no discernible effect of the worked example on students' perceptions at any stage of the research, but did find statistically significant differences between the treatment and control groups on the majority of measured student behaviors in their subsequent online debate. This finding contrasts from the majority of related literature those strongly related student perceptions of readiness and satisfaction to performance. These findings would seem to indicate that if the groups in the current study felt equally well-prepared, equally satisfied with the instructional activity, and equally in control of their navigation and ability to learn, their subsequent performance should have been substantively equivalent.

As previously discussed, the majority of the previous research that informed this study measured the students perceived need for preparation without the aid of any specific preparedness activity. The present study and the study by Rummel et al. (2009) were the only examples of research informing this study that measured student perception against the backdrop of specific ways to prepare the learner. In both cases, a worked example was used and in both cases, no statistically significant differences in student perception were found between students being prepared through a worked example and those receiving no prior instruction. It would be worthwhile to replicate these studies to determine if these results are anomalous or if, in fact, it is reasonable to expect that worked examples impact student behaviors without necessarily improving or changing their underlying perception of the preparation or the activity itself.

Although the areas of student perception in the current study were consistently identified in the previous research, future research might also include additional perception variables. Such variables should be measured without diminishing the importance of the *preparedness*, *productive learning environment*, and *learner control* perception domains. These measures should shed light on the impact of worked example on elements such as student motivation, self-direction, learning style, etc. Future research might also examine student perceptions about specific behaviors such as indicating confidence in criticizing the arguments of others, posting information in support of an argument, or understanding the rigid structure of the online debate.

Another area warranting future research is the impact of the characteristics of the student population on the influence of the worked examples. There is evidence found in the research that many characteristics of the population used in the present study held the possibility of affecting the results. Fogerson (2005), for instance, found that self-direction and age were positively correlated. In the present study, 77% of the students were older than traditionally-aged college students meaning that these students may have been less likely to need the direction of a worked example. Gender was also shown to play a role in how students in how students participated in online discussion forums of all types. Female students have been shown to use more qualifiers (Jeong, 2005b), make shorter postings (Barrett & Lally, 1999; Herring, 2000), post less frequently (Herring, 1993; Jeong & Davdison-Shivers, 2003), and use fewer argumentative exchanges (Jeong & Davidson-Shivers). Eighty percent (80%) of the students in this study were female and it is possible that the viability of the worked example in promoting these behaviors may have been affected. Fogerson (2005) also found a relationship between prior learning experience and student satisfaction. Ninety percent (90%) of the students in the current study had prior learning experience indicating they may have had a predisposition to be satisfied with the online debate. Future research should be conducted with student populations with more traditional-aged students, a more balanced gender composition, and a larger number of students novice to online learning.

Thirdly, further study is needed in determining if different types of worked example hold more potential in preparing learners. Rummel et al. (2009) created four different conditions (two of which were clear demonstrations of worked examples) with different levels of support and elaboration. In a similar vein to Rummel et al., research into different applications of worked example should continue. Worked examples might also be manipulated to excessively model a desired behavior to determine is such modeling will promote that behavior. For example, a future research might present a worked example with an exorbitantly high number of linguistic intensifiers to determine if students exposed to the example would, in turn, use more intensifiers. Additionally, worked examples should be applied to different instructional formats. Rummel et al. found that worked examples influenced student behavior within a two-person collaborative problem-solving exercise and this study reached the same conclusion for highly-structured argumentative exercises. More research should be performed to ensure that worked examples are as effective for other instructional formats (e.g. WebQuests, unstructured exercises, threaded discussions).

Lastly, it would be prudent for future attempts at research to use different measurable behaviors. The selection of measured behaviors for the current study was subjective and based on the research questions that drove it and the research that informed it. Other behaviors worth considering in future research would be metacognition (Hara et al., 2000; Henri, 1992), depth of information processing (Henri), discussion spent on task (Hirsch et al., 2004; Jeong & Joung, 2004; Jonassen & Redmidez, 2005), and group versus task focus (Howell-Richardson & Mellar, 1996).

Conclusions

This study was intended to further the understanding of better preparing learners for participation within interactive exercises in online courses. The present study built on many areas popular in contemporary distance learning research including improving course design, increasing learner control and reducing transactional distance, and more generally improving the quality of interaction in online courses. This study extended upon previous research on two key fronts. First, student perception and student learning behaviors were measured while virtually all previous research endeavors investigated one aspect or the other. Secondly, student perceptions and learning were examined within the context of the effects of a specific instructional intervention. In this study, that intervention was a worked example.

The results of this study indicated that preparing students through a worked example did, in fact, hold the potential to improve the quality of the student's interaction if not necessarily their perceptions about the interaction. Additional research is needed to not only substantiate that worked examples can be used to improve the quality of online debates, but also to investigate the possibility of using worked examples and other techniques to better prepare learners for a wide range of interactive exercises. If learner preparedness studies continue to be performed and reach similar conclusions, the possibility exists that specific preparation techniques will be linked with improving the quality of a large number of interactive online activities.

REFERENCES

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

Barbera, E. (2006). Collaborative knowledge construction in highly structured virtual discussions. *The Quarterly Review of Distance Education*, *7*(1), 1-12.

Barrett, E., & Lally, V. (1999). Gender differences in an on-line learning environment. *Journal of Computer Assisted Learning*, *15*, 48-60.

Berge, Z., & Mrozowski, S. (2001) Review of research in distance education, 1990-1999: *The American Journal of Distance Education*, *15*(3), 5-19.

Blignaut, A., & Trollip, S. (2003). Measuring faculty participation in asynchronous discussion forums. *Journal of Education for Business, 78*(6), 347-353.

Bocchi, J., Eastman, J.K., & Swift, C.O. (2004). Retaining the online learner: Profile of students in an online MBA program and implications for teaching them. *Journal of Education for Business, 79(*4*),* 245-253.

Bozarth, J., Chapman, D., & LaMonica, L. (2004). Preparing for distance learning: designing an online student orientation course. *Educational Technology & Society, 7*(1), 87-106.

Carswell, L., Thomas, P., Petre, M., Price, B., & Richards, M. (2000). Distance education via the Internet: The student experience. *British Journal of Educational Technology*, *31*(1), 29-46.

Chen, C. (2003). Managing perceptions of information overload in computer-mediated communication. Unpublished doctoral dissertation, Texas A&M University.

Cho, K., & Jonassen, D. (2002). The effects of argumentation scaffolds on argumentation and problem solving. *Educational Technology Research and Development*, *50*(3), 5-22.

Clark, R. (1994). Media will never influence learning. *Educational Technology Research and Development, 42*(2), 21-29.

Clark, R. (2003). Research on web-based learning: A half-full glass. In R. Bruning, C. Horn, & L. PytlikZillig (Eds.), *Web-based learning: What do we know? Where do we go?*, (pp. 1-22). Greenwich, CT: Information Age Publishing.

Conrad, D. (2002). Engagement, excitement, anxiety, and fear: Learners' experiences of starting an online course. *The American Journal of Distance of Education, 16*(4)*,* 205-226.

de Bruyn, L. (2004). Monitoring online communication: Can the development of convergence and social presence indicate an interactive learning environment? *Distance Education*, *25*(1), 67-91.

de Wever, B., Schellens, M., Valcke, M., & Van Keer, H. (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups. A review. *Computers in Education*, *46*, 6-28.

Fahy, P. (2001). Addressing some common problems in transcript analysis. *The International Review of Research in Open and Distance Learning*, *1*(2). Retrieved June 12, 2007, from http://www.irrodl.org/index.php/irrodl/ article/viewArticle/321/530

Fahy, P. (2002). Use of linguistic qualifiers and intensifiers in a computer conference. *The American Journal of Distance Education*, *16*(1), 5-22.

Fahy, P., Crawford, G., & Ally, M. (2001). Patterns of interaction in a computer conference transcript. *International Review of Research in Open and Distance Learning*, *2*(1). Retrieved December 22, 2007, from http://www.icaap.org/iuicode?149.2.1.4

Fogerson, D. (2005). Readiness factors contributing to participant satisfaction in online higher education courses. Unpublished doctoral dissertation, University of Tennessee, Knoxville.

Fraenkel, J., & Wallen, N. (2006). *How to design and evaluate research in education, (6th ed)*. New York: McGraw-Hill.

Gorsky, P., & Caspi, A. (2005). A critical analysis of transactional distance theory. *The Quarterly Review of Distance Education*, *6*(1), 1-11.

Gravetter, F., & Wallnau, L. (2004). *Statistics for the behavioral sciences (6th ed.)*. Belmont, CA: Wadsworth/Thomson Learning.

Gunawardena, C., Lowe, C., & Anderson, T. (1998). Transcript analysis of computer-mediated conferences as a tool for testing constructivist and social-constructivist learning theories. Madison, WI: *Distance Learning '98 Proceedings of the Annual Conference on Distance Teaching and Learning*. (ERIC Document Reproduction Service No. ED422854)

Hair, J., Black, W., Babin, B., Anderson, R., & Tatham, R. (2006). *Multivariate data analysis (6th ed)*. Upper Saddle River, NJ: Rearson.

Hannafin, M., Land, S., & Oliver, K. (1999). Open learning environments: Foundations, methods, and models. In C.M. Reigeluth (Ed.), *Instructional-design theories and models* (Vol. II). Mahwah, NJ: Erlbaum.

Hara, N., Bonk, C., & Angeli, C. (2000). Content analysis of online discussion in an applied educational psychology course. *Instructional Science*, *28*(2), 115-152.

Henri, F. (1992). Computer conferencing and content analysis. In A. Kaye (Ed.), *Collaborative learning through computer conferencing: The Najaden papers*, (pp. 115–136). New York: Springer.

Herring, S. (1994, June). *Gender differences in computer-mediated communication: Bringing familiar baggage to the new frontier*. Paper presented at the annual convention of the American Library Association, Miami, FL.

Herring, S. (2000). Gender differences in CMC: Findings and implications. *The CPSR Newsletter*, *18*(1). Retreived April 18, 2006, from http://cpsr.org/issues/womenintech/herring/

Hirsch, L., Saeedi, M., Cornillon, J., & Litosseliti, L. (2004). A structured dialogue tool for argumentative learning. *Journal of Computer Assisted Learning, 20*(1)*,* 72-80.

Hodgkinson-Williams, C., & Mostert, M. (2005). Online debating to encourage student participation in online learning environments: A qualitative case study at a South African university. *International Journal of Education and Development using ICT* [Online], *1*(2), 94-104.

Howell-Richardson, C., & Mellar, H. (1996). A methodology for the analysis of patterns of participation within computer mediated communication courses. *Instructional Science, 24*(1), 47-69.

Hron, A., & Freidrich, H. (2003). A review of web-based collaborative learning: Factors beyond technology. *Journal of Computer Assisted Learning*, *19*(1), 70-79.

Jeong, A. (2003). The sequential analysis of group interaction and critical thinking in online threaded discussions. *American Journal of Distance Education*, *17*(1), 25-43.

Jeong, A. (2004). The combined effects of response time and message content on growth patterns of discussion threads in computer-supported collaborative argumentation. *Journal of Distance Education*, *19*(1), 36-53.

Jeong, A. (2005a). *Discussion analysis tool*. Retrieved on April 21, 2006 from http://garnet.fsu.edu/~ajeong/DAT

Jeong, A. (2005b). The effects of linguistic qualifiers on group interaction patterns in computer-supported collaborative argumentation. *International Review of Research in Open and Distance Learning*, *6*(3). Retrieved March 16, 2006, from http://www.irrodl.org/index.php/irrodl/article/view/258/401

Jeong, A. (2005c). A guide to analyzing message-response sequences and group interaction in computer-mediated communication. *Distance Education*, *26*(3), 367-383.

Jeong, A. (2006). Gender interaction patterns and gender participation in computer-supported collaborative argumentation. *American Journal of Distance Education*, *20*(4), 195-210.

Jeong, A., & Davidson-Shivers, G. (2003, April). *Gender interactions in online debates: Look who's arguing with whom*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL. Retrieved March 16, 2006, from http://dev2244801.sp01.fsu.edu/Research/GenderInteraction/GenderInteraction.pdf

Jeong, A., & Joung, S. (2004). *The effects of response constraints and message labels on interaction patterns and argumentation in online discussions*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA. Retrieved September 26, 2005, from http://dev2244801.sp01.fsu.edu/Research/Labeling/Fall2002/AERAProposalScaffolding.pdf

Jeong, A., & Joung, S. (2007). Scaffolding collaborative argumentation in asynchronous discussions with message constraints and message labels. *Computers & Education*, *48*(3), 427-445.

Jonassen, D., & Remidez, H. (2005). Mapping alternative discourse structures onto computer conferences. *International Journal of Knowledge and Learning*, *1*(1/2), 113-129.

Joung, S., & Keller, J. (2004, October). *The effects of high-structure cooperative versus low-structure collaborative design of design change, critical thinking, and interaction pattern during online debates*. Paper presented at the international convention of the Association for Educational Communications and Technology, Chicago, IL. (ERIC Document Reproduction Service No. ED485151)

Kanuka, H.,Collett, D., & Caswell, C. (2002). University instructor perceptions of the use of asynchronous text-based discussion in distance courses. *American Journal of Distance Education*, *16*(3), 151-167.

Kanuka, H., Rourke, L., & Laflamme, E. (2007). The influence of instructional methods on the quality of online discussion. *British Journal of Educational Technology*, *38*(2), 260-271.

Kawachi, P. (2003). Initiating intrinsic motivation in online education: Review of the current state of the art. *Interactive Learning Environments, 11*(1), 59-81.

King, T., & Duke-Williams, E. (2001). *Using computer aided assessment to test higher learning outcomes*. Paper presented at the 5th Computer-Assisted Assessment Conference, Loughborough, UK.

Krippendorff, K. (2004). *Content analysis. An introduction to its methodology*. Thousand Oaks, CA: Sage Publications.

Kreijns, K., Kirschner, P. A., & Jochems, W. (2002). The sociability of computer-supported collaborative learning environments. *Journal of Education Technology & Society*, *5*(1), 8-22.

LaPadula, M. (2003). A comprehensive look at online student support services for distance learners. *American Journal of Distance Education, 17*(2), 119-128.

Lee, Y., Driscoll, M., & Nelson, D. (2004). The past, present, and future of research in distance education: Results of a content analysis. *The American Journal of Distance Education*, *18*(4), 225-241.

Li, R. (2005). The effects of databases as cognitive tools in a multimedia problem-based learning environment. Unpublished doctoral dissertation, University of Texas, Austin.

Lockee, B., Burton, J., & Cross, L. (1999, February). *No comparison: Distance education finds a new use for 'no significant difference'*. Paper presented at the national convention of the Association for Educational Communications and Technology, Chicago, IL. (ERIC Document Reproduction Service No. ED436176)

Marra, R., Moore, J., & Klimczak, A. (2004). Content analysis of online discussion forums: A comparative analysis of protocols. *Educational Technology Research and Development*, *52*(2), 23-40.

McIsaac, M., & Blocher, J. (1998). How research in distance education can affect practice. *Educational Media International, 35*(1), 43-47.

McKlin, T., Harmon, S., Evans, W., & Jones. M. (2002). Cognitive presence in Web-based learning: A content analysis of students online discussions. *ITFORUM*, *60*. Retrieved December 27, 2007, from http://it.coe.uga.edu/itforum/paper60 /paper60.htm

Menlove, R. and Lignugaris, B. (2004). Preparing rural distance education preservice special educators to succeed. *Rural Special Education Quarterly, 23*(2)*,* 18-26.

Mertler, C., & Vannatta, R. (2005). *Advanced and multivariate statistical methods. Practical application and interpretation (3rd ed.).* Glendale, CA: Pryczak Publishing.

Meyer, K. (2002). The significance of 'no significant difference'. In *Quality in Distance Education: ASHE-ERIC Higher Education Report* (29th ed., pp. 13-20). San Francisco, Josey-Bass.

Moore, M., & Kearsley, G. (2005). *Distance education: A systems view (2nd ed.)*. Belmont, CA: Wadsworth.

Muilenberg, L., & Berge, Z. (2000). A framework for designing questions for online learning. *DEOSNEWS*, *10*(2). Retrieved from http://www.ed.psu.edu/acsde/ deos/deosnews/deosnews10_2.asp

Muirhead, B. (2002). Promoting online interaction in today's universities and colleges. *USDLA Journal*, *16*(7), 43-47. Retrieved March 12, 2006, from http://www.usdla.org/html/journal/JUL02_Issue/article04.html

National Center for Educational Statistics (2009). Distance education at degree-granting postsecondary institutions: 2006-2007. Retrieved April 17, 2009, from http://nces.ed.gov/pubs2009/2009044.pdf

Neuendorf, K. (2002). *The content analysis handbook*. Thousand Oaks, CA: Sage Publications.

Northrup, P. (2002). Online learners' preferences for interaction. *The Quarterly Review of Distance Education, 3*(2), 219-226.

Oh, S., & Jonassen, D. (2007). Scaffolding online argumentation during problem solving. *Journal of Computer Assisted Learning*, *23*(2), 95-110.

Palloff, R., & Pratt, K. (2003). *The virtual student. A profile and guide to working with online learners*. San Francisco: Jossey-Bass.

Pedersen, S., & Liu, M. (2002). The effects of modeling expert cognitive strategies during problem-based learning. *Journal of Educational Computing Research*, *26*(4), 353-380.

Perreault, H., Waldman, L., Alexander, M., & Zhao, J. (2002). Overcoming barriers to successful delivery of distance-learning courses. *Journal of Education for Business,77*(6)*,* 313-318.

Reigeluth, C., & Moore, J. (1998). Cognitive education and the cognitive domain. In C. Reigeluth (Ed.), *Instructional design and theories and models*, (pp. 51-68). Mahwah, NJ: Lawrence Erlbaum Associates.

Reisetter, M., & Boris, G. (2004). What works: Student perceptions of effective elements in online learning. *The Quarterly Review of Distance Education, 5*(4), 277-291.

Resta, P., & Laferriere, T. (2007). Technology in support of collaborative learning. *Educational Psychology Review*, *19*(1), 65-83.

Rourke, L., & Anderson, T. (2004). Validity in quantitative content analysis. *Educational Technology Research and Development, 52*(1), 5-18.

Rourke, L., Anderson, T., Garrison, D., & Archer, W. (2001). Methodological issues in the content analysis of computer conference transcripts. *International Journal of Artificial Intelligence in Education*, *12,* 8-22.

Rummel, N., Spada, H., & Hauser, S. (2009). Learning to collaborate while being scripted or by observing a model. *International Journal of Computer-Supported Collaborative Learning*, *4*(1), 69-92.

Russell, T. (1999). *The no significant difference phenomenon: A comparative research annotated bibliography on technology for distance education*. Raleigh, NC: Office of Instructional Communications, North Carolina State University.

Sam, C. (2005, June). *Purposeful scaffolding: Beyond modeling and thinking aloud*. Paper presented at the Redesigning Pedagogy Conference, Singapore. Retrieved September 22, 2007, from http://www.elearnjourney.com/Converted%20Pdf/ab00655.pdf

Scheines, R., Leinhardt, G., Smith, J., & Cho, K. (2005). Replacing lecture with web-based course materials. *Journal of Educational Computing Research*, *32*(1), 1-26.

Scheiter, K. & Gerjets, P. (2007). Learner control in hypermedia environments. *Educational Psychology Review*, *19*(3), 285-307.

Schrire, S. (2006). Knowledge building in asynchronous discussion groups: Going beyond quantitative analysis. *Computers & Education*, *46*(1), 49-70.

Shin, N. (2003). Transactional presence as a critical predictor of success in distance learning. *Distance Education, 24*(1), 69-86.

Song, H. (2003). Development of a systematic assessment framework for analyzing interaction in online environments. *The Quarterly Review of Distance Education, 4*(4), 437-444.

Sorensen, C., & Baylen, D. (2004). Patterns of communicative and interactive behavior online: Case studies in higher education. *The Quarterly Review of Distance Education, 5*(2), 117-126.

Sweller, J., Merrienboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Education Psychology Review*, *10*(3), 254-296.

Vamosi, A., Pierce, B., & Slotkin, M. (2004). Distance learning in an accounting principles course – student satisfaction and perceptions of efficacy. *Journal of Education for Business, 79*(6)*,*360-366.

Van Gog, T., Paas, F., & van Merrienboer, J. (2004). Process-oriented worked example: Improving transfer performance through enhanced understanding. *Instructional Science*, *32*(1-2), 83-98.

Vrasidas, C., & Glass, G. (2002). A conceptual framework for studying distance education. In C. Vrasidas & G. Glass (Eds.), *Distance education and distributed learning* (pp. 31-55). Greenwich, CT: Information Age Publishing.

Wickersham, L., & Dooley, K. (2006). A content analysis of critical thinking skills as an indicator of quality of online discussion in virtual learning communities. *Quarterly Review of Distance Education*, *7*(2), 185-193.

Wong-Bushby, I., Passerini, K., Hiltz, S., Bieber, M., & Patten, K. (2005, August). Scaffolding discourse in asynchronous learning networks. Omaha, NE: *Proceedings of the Eleventh Americas Conference on Information Systems.* Retrieved September 22, 2007, from http://afs1.njit.edu/~bieber/pub/wb-ais2005.pdf

APPENDIX A

INSTITUTIONAL REVIEW BOARD APPROVAL

Mississippi State UNIVERSITY

April 7, 2008

Scott Tollison
Mississippi University for Women
College of Business
W-Box 940

RE: IRB Study #08-103: The Influence of Preparedness on Student Perceptions and the Quality of Online Discussion

Dear Tollison:

The above referenced project was reviewed and approved via administrative review on 4/7/2008 in accordance with 45 CFR 46.101(b)(2). Continuing review is not necessary for this project. However, any modification to the project must be reviewed and approved by the IRB prior to implementation. Any failure to adhere to the approved protocol could result in suspension or termination of your project. The IRB reserves the right, at anytime during the project period, to observe you and the additional researchers on this project.

Please refer to your IRB number (#08-103) when contacting our office regarding this application.

Thank you for your cooperation and good luck to you in conducting this research project. If you have questions or concerns, please contact irb@research.msstate.edu or 325-3294.

Sincerely,

Katherine Crowley
Assistant IRB Compliance Administrator

cc: Dr. Kui Xie

APPENDIX B

INFORMED CONSENT AGREEMENT

Introduction
I am a doctoral student at Mississippi State University and would like to include you in a research study. The purpose of this study is to examine the frequency and types of participation in online discussions.

Participation
Complete short periodic online questionnaires concerning your feelings and participation in online discussion. These questionnaires should require a total effort of between 5-10 minutes throughout the semester. Election not to participate in this study will not negatively effect your grade in the class.

Risks and Benefits
The researcher foresees no anticipated risk or discomfort to those who choose to participate in this study. This survey offers no direct benefit to you and you will not be paid for your participation. However, your participation in this study will assist the researcher in adding to a body of work that aims to improve the quality of distance education at not only this University, but other institutions as well.

Confidentiality
The results of this study may be published but your name or identity will not be revealed. The online questionnaires will not be viewed on an individual basis, but instead on the basis of the aggregate of the class. Therefore, no student should fear the consequences of having his or her responses monitored.

Contact
Any questions you have concerning the study or your participation in it, before or after this consent, may be answered by:
Scott Tollison
Office: Room 301, Reneau Hall
Phone: (662) 329-7164
E-mail: stollison@muw.edu

Agreement
I have read this informed consent agreement form and am above 18 years of age. I understand that I may withdraw my consent and discontinue participation at any time without penalty or loss of benefits to which I may otherwise be entitled. In typing my full name into the textbox below, I am not waiving any legal claims, rights, or remedies. A copy of this consent form will be offered to me upon my request.

Please type your name into the textbox below.

Answer:

APPENDIX C

STUDENT PERCEPTION INSTRUMENT

Question 1
What is your gender?
a. Male
b. Female

Question 2
What is your age?
a. 18-23
b. 24-29
c. 30-35
d. 36 or above

Question 3
Which best describes your ethnicity?
a. African American
b. Asian American
c. Hispanic American
d. Native American
e. White/Caucasian American
f. Other

Question 4
What is your student classification?
a. Freshman
b. Sophomore
c. Junior
d. Senior

Question 5
How many online courses have you previously taken?
a. None
b. One
c. Two
d. More than two

Question 6
Have you personally met your professor before?
a. Yes
b. No

Question 7
Do you take any courses on-campus in a traditional classroom?
a. Yes
b. No

Question 8
I have read the tutorial provided by my instructor. If not, please read the tutorial before answering the remaining questions.
a. Yes
b. No

Please indicate how strongly you agree or disagree with the following statements.

Question 9
The instructor has clearly stated what is expected of me in this course.
5. Strongly Agree
4. Somewhat Agree
3. Neutral
2. Somewhat Disagree
1. Strongly Disagree

Question 10
I am enthusiastic about this online course.
5. Strongly Agree
4. Somewhat Agree
3. Neutral
2. Somewhat Disagree
1. Strongly Disagree

Question 11
After reviewing the tutorial, I feel confused about the online debate procedures.
5. Strongly Agree
4. Somewhat Agree
3. Neutral
2. Somewhat Disagree
1. Strongly Disagree

Question 12
After reviewing the tutorial provided to me by my instructor, I feel in control of my learning within the online debate.
5. Strongly Agree
4. Somewhat Agree
3. Neutral
2. Somewhat Disagree
1. Strongly Disagree

Question 13
After reviewing the tutorial provided to me by my instructor, I feel as though I will be comfortable navigating through the postings within an online debate.
5. Strongly Agree
4. Somewhat Agree
3. Neutral
2. Somewhat Disagree
1. Strongly Disagree

Question 14
I am confused about what is expected of me in this course.
5. Strongly Agree
4. Somewhat Agree
3. Neutral
2. Somewhat Disagree
1. Strongly Disagree

Question 15
After reviewing the tutorial, I feel as though I have been provided enough information about the online debate procedures to be successful.
5. Strongly Agree
4. Somewhat Agree
3. Neutral
2. Somewhat Disagree
1. Strongly Disagree

Question 16
After reviewing the tutorial, I feel adequately prepared to participate in the online debate.
5. Strongly Agree
4. Somewhat Agree
3. Neutral
2. Somewhat Disagree
1. Strongly Disagree

Question 17
I am apprehensive about this online course.
5. Strongly Agree
4. Somewhat Agree
3. Neutral
2. Somewhat Disagree
1. Strongly Disagree

Question 18
After reviewing the tutorial, I feel confident about my successful participation in the online debate.

5. Strongly Agree
4. Somewhat Agree
3. Neutral
2. Somewhat Disagree
1. Strongly Disagree

Question 19
I am confident about successfully completing the requirements of this online course.

5. Strongly Agree
4. Somewhat Agree
3. Neutral
2. Somewhat Disagree
1. Strongly Disagree

Question 20
After reviewing the demonstration of an online debate provided to me by my instructor, I feel as though I have been provided enough information about the online debate procedures to be successful.

5. Strongly Agree
4. Somewhat Agree
3. Neutral
2. Somewhat Disagree
1. Strongly Disagree

Question 21
After reviewing the tutorial provided to me by my instructor, I feel as though I will be satisfied with the online debate experience.

5. Strongly Agree
4. Somewhat Agree
3. Neutral
2. Somewhat Disagree
1. Strongly Disagree

Question 22
After reviewing the tutorial, I feel unprepared to participate in the online debate.

5. Strongly Agree
4. Somewhat Agree
3. Neutral
2. Somewhat Disagree
1. Strongly Disagree

Question 23
After reviewing the tutorial, I am eager to participate in the online debate.
5. Strongly Agree
4. Somewhat Agree
3. Neutral
2. Somewhat Disagree
1. Strongly Disagree

Question 24
After reviewing the tutorial provided to me by my instructor, I feel as though the online debates will hurt my learning.
5. Strongly Agree
4. Somewhat Agree
3. Neutral
2. Somewhat Disagree
1. Strongly Disagree

Question 25
After reviewing the tutorial provided to me by my instructor, I feel as though my learning within the debate is out of my control.
5. Strongly Agree
4. Somewhat Agree
3. Neutral
2. Somewhat Disagree
1. Strongly Disagree

Question 26
After reviewing the tutorial provided to me by my instructor, I feel as though I might be uncomfortable navigating through the postings within an online debate.
5. Strongly Agree
4. Somewhat Agree
3. Neutral
2. Somewhat Disagree
1. Strongly Disagree

Question 27
After reviewing the tutorial provided to me by my instructor, I feel anxiety about the online debate experience.
5. Strongly Agree
4. Somewhat Agree
3. Neutral
2. Somewhat Disagree
1. Strongly Disagree

Question 28
After reviewing the tutorial provided to me by my instructor, I feel as though the online debate will further my learning.

5. Strongly Agree
4. Somewhat Agree
3. Neutral
2. Somewhat Disagree
1. Strongly Disagree

APPENDIX D

ONLINE DEBATE TUTORIAL

During the course we will have the opportunity to participate in an online debate. In this activity, you will be assigned to a team to either **support** or **oppose** the debate topic/statement/question. Through this debate, you will explore the differences in recommendations as well as debate alternative solutions. This debate is a unique opportunity to interact with your peers and to compare your perspectives with those of other participants in this course. It is also an opportunity for you to apply the concepts you discover in your research to support your position.

You will be assigned to one of two teams to debate the topic/statement/question. You **must** only argue for the position you are assigned to defend. Teams will also need to research and find sources to support their team position for the debate. Discussion areas will be created for each team, so that they can prepare for the debate and share information.

During the debate, refer to the research you have gathered to find evidence to argue for or against the assumption or to find perspectives from which to critique the arguments of the other team.

Rules & Protocols: There are a set of instructions you must follow. **Follow the instructions carefully! Failure to follow the procedures will result in zero points for this assignment**.

Here are your rules for posting.

Subject Heading Tags	
ARG	A main argument or assertion to support a position. It should be no longer than one sentence.
EVID	To *support* argument & assertions with *evidence*, examples, studies, personal experiences. It generally should include a link to an article.
CRIT	Critique, test/question validity, request supporting evidence, identify flaw in argument, logic, evidence
QUES	Ask question *only* for clarification (*not* to challenge or critique).

Insert one and only one tag into the subject heading of each message (use ALLCAPS). *Immediately* following the tag, add s = supporting or o = opposing to identify team membership (eg. ARGs and ARGo). Include a short but meaningful message title (eg. ARGs my message title). Address only one function per message.

You will be restricted to posting specific types of messages in the debate. Each message you post must be labeled to identify its function, and each message must address only *one* function at a time. The purpose of the labels is to improve the organization, quality, and structure of the discussions. For example, you can use the WebCT search tool to find how many arguments have been posted by the opposing teams (e.g. ARGo vs ARGs). Scanning the labels in the discussion forum will help you see the flow and structure of the discussions.

Examples: Here are some examples of using some of the tags.

For example if I was going to post an argument that *opposes* the use of the color blue on classroom walls. My subject line might be:

ARGo The color blue causes individuals to become drowsy

Then in the body of my posting, I would post an argument with evidence that associates the color blue with drowsiness.

Or if I wanted to post an argument that *supports* the use of color blue on classroom walls. My subject line might be:

ARGs The color blue creates a sense of energy for students in the room.

Then in the body of my posting, I would post an argument with evidence that associates the color blue with energy.

If someone challenged this or if I wanted to post an article to oppose this, then I might reply to my own statement with a posting with the following subject line:

EVIDo Jones (2003) found that there was a relationship between blue walls & falling asleep in class

Then in the body of my posting, I would explain or summarize the article as well as include an attachment or a link to the article that supports the earlier argument or I would list the A.P.A. reference to the article.

If I wanted to counter an opponent's argument or find fault with an opponent's argument, then I might reply with the following subject line:

CRITo Color has no effect on classroom performance

Then in the body of my posting, I would make my case about color not affecting classroom performance. Similarly, you might also consider the CRIT tag in response to the arguments of your opponent. CRIT tags require a greater burden of proof in criticizing the opponent's arguments.

Discussion Area: Information about your group's discussion area.

As mentioned previously, to give your team an opportunity to prepare for the debate discussion areas have been created for each team and each team has a chat room available and labeled for each team if they choose to use them. You will need to view this discussion area like you would a group meeting – an opportunity to interact with one another and share ideas as well as to plan and prepare a strategy for winning the debate. After all, your goal is to WIN the debate and humiliate your opponents in the process (just joking about the humiliation part).

The group discussion areas are private. The only persons who will see these areas are your teammates and the instructor. Therefore, you can share information and prepare strategies outside the view of your opponent.

General Information: Miscellaneous information about the debate

Participating in the online debate will require much greater effort than our weekly online discussions. It will require logging into WebCT on a daily basis to monitor the debate. A single posting for the week will not suffice. It will require making multiple postings during the course of the week in order to support your team's arguments as well as to counter those arguments made by your opponents. Constant vigilence is a price that will have to be paid to be successful in your debate.

The teams were randomly assigned by Microsoft Excel. You may not switch teams under any circumstance.

Unlike the weekly discussions which are graded largely on the basis of effort, the debates will be evaluated on a much more stringent scale. Your grade will be determined on a variety of factors including frequency of postings, length of postings, quality of postings, evidence of research, and adherence to debate rules and protocols. A rubric will be provided that demonstrates how the debates will be scored.

Remember to be civil and polite in your postings. We are all friends here and our goal is to maximize our understanding of information systems as well as hone our ability to articulate our ideas in writing. We are not here to tear anyone down or hurt anyone's feelings. I reserve the right to remove any postings that violate the spirit of debate or are hostile and inflammatory in nature.

Helpful Hints: Here are some things to focus on.

When posting messages in response to the postings of others, be sure to press the "Reply" button. If posting a totally new argument or position, you need to press the "Compose Message" button.

When viewing your debate on the discussion board, it is much easier to get a feel for the structure of the debate if you ensure that the "All" and "Threaded Button" are selected.

When using online articles or Web sites as evidence, please be sure to copy and paste the link into the discussion postings.

APPENDIX E

ONLINE DEBATE RUBRIC

	5 Excellent	4 Very Good	3 Average	2 Below Average	1 Poor	Score
Frequency of Postings	Student made postings far in excess of average student.	Student made postings noticeably greater than average student.	Student made postings similar to that of average student.	Student made noticeably fewer postings than average student.	Student made no or almost no postings.	____/5
Quality of Postings	Student showed significant ability in articulating his or her position.	Student articulated his or her position well.	Student articulated his or her position adequately.	Student articulated his or her position less than adequately.	Student articulated his or her position poorly.	____/5
Adherence to Rules	Student followed debate rules in virtually every posting.	Student followed debate rules in most posting.	Student followed debate rules in more than half of the postings.	Student followed debate rules in less than half of the postings.	Student followed debate rules in virtually no postings.	____/5
Evidence of Research	Student demonstrated significant outside research. Used linked articles in most postings.	Student demonstrated good outside research. Used linked articles where appropriate.	Student demonstrated some outside research. Seldom used linked articles.	Student demonstrated little outside research. Virtually no use of linked articles.	Student demonstrated no outside research. No use of linked articles.	____/5
Teammate Feedback	Student's performance rated "excellent" by peers.	Student's performance rated "very good" by peers.	Student's performance rated "average" by peers.	Student's performance rated "below average" by peers.	Student's performance rated "poor" by peers.	____/5
					Total:	____/25