ICIS 2005 Proceedings

International Conference on Information Systems (ICIS)

December 2005

Virtual Interaction for Effective E-Learning

Jinwei Cao
University of Delaware

Ming Lin
University of Arizona

Janna Crews
University of Nevada, Reno

Judee Burgoon
University of Arizona

Jay Nunamaker, Jr.
University of Arizona

VIRTUAL INTERACTION FOR EFFECTIVE E-LEARNING

Jinwei Cao
University of Delaware
Newark, DE U.S.A.
jcao@lerner.udel.edu

Ming Lin
University of Arizona
Tucson, AZ U.S.A.
mlin@email.arizona.edu

Janna M. Crews
University of Nevada, Reno
Reno, NV U.S.A.
jcrews@unr.edu

Judee K. Burgoon
University of Arizona
Tucson, AZ U.S.A.
jburgoon@cmi.arizona.edu

Jay F. Nunamaker, Jr.
University of Arizona
Tucson, AZ U.S.A.
nunamaker@cmi.arizona.edu

Abstract

This research investigates whether information technologies, such as automated question answering (QA), can add interactivity into a multimedia-based e-learning system, as well as how this type of virtual interaction affects the effectiveness of e-learning. Based on a review of multiple learning theories and technologies, an exploratory model for studying the effectiveness of interactive e-learning, Learning with Virtual Mentors (LVM), is proposed and a prototype system is developed to implement the LVM model. A series of studies, including a controlled experiment and surveys, have been conducted to explore the relationships among the core constructs of the LVM model: learning phases, system interactivity, learning activity, and learning outcomes. Findings indicate that virtual interaction positively impacts student behaviors by encouraging students to interact more and increasing student satisfaction with the learning process; however, the correlation between virtual interaction and actual learning performance is limited. Consequently, the LVM model needs to be further explored and developed.

Keywords: Virtual interaction, e-learning, question answering, effectiveness

Introduction

Learning is critical for any individual or organization to be successful in the current knowledge-based economy. In recent years, advances in information technology have affected the learning market dramatically. Thousands of online courses, including degree and certificate programs, are now offered by universities and corporations worldwide. Most of these technology supported learning or training programs are grouped under the *de facto* term *e-learning*. The fundamental value proposition of e-learning—access to quality education and training freed from the boundaries of time and location—is growing with the demand for higher education and professional training in the United States and worldwide.

In current e-learning programs, multimedia learning materials such as videotaped lectures and PowerPoint slides are increasingly provided as a way to help learners engage in the learning process. However, simply watching an instructor lecturing in a video is different from learning with a human mentor. An important factor of learning, *interaction*, is usually missing in multimedia online lectures (O'Connor et al. 2003). Many modern learning theories indicate that effective learning requires an iterative interaction process between the learner and the knowledge providers (Bruner 1960, 1966; Pask 1975); however, such a process cannot be supported solely by a linear playback of the video lectures.

Although collaborative learning technologies such as chat rooms and discussion forums can be used to provide a platform for interaction in e-learning, such interaction relies on the availability of human mentors. When a learner has a question regarding the learning content and needs an immediate answer (e.g., for an immediate task required on the job), the human mentor may not be accessible at that specific time. On the other hand, it is tedious and time-consuming for learners to search a 60-minute-long, linear, unstructured video for answers to their specific questions. Furthermore, a Web search for the desired information may not only take a lot of time, but could result in questionable answers rather than answers obtained from the mentor. This research endeavors to address these issues.

Specifically, this research investigates whether information technologies, such as automated question answering (QA), can add interactivity into a multimedia-based e-learning system and how this type of "virtual interaction" affects the effectiveness of e-learning. Based on a review of multiple learning theories and technologies in the next section of this paper, an exploratory model for studying the effectiveness of interactive e-learning, named Learning with Virtual Mentors (LVM), is proposed. The LVM model and a prototype system developed to implement the LVM model are described in the third section. A series of studies, including a controlled experiment and surveys have been conducted to explore the relationships among the core constructs of the LVM model: learning phases, system interactivity, learning activity and learning outcomes. The core studies in this research are discussed in the subsequent sections.

Research Background

A Framework for Technology Mediated Learning Research

Research about using information technology to support learning (technology mediated learning, TML) started as early as the 1960s; however, most of the previous studies focused on the direct influence of technology features on learning outcomes. A guideline for theoretically grounded and rigorous research of TML is missing. In 2001, Alavi and Leidner (2001) proposed a framework for TML research, which is illustrated in Figure 1. This framework suggests that TML research should consider all aspects of instructional strategies, information technologies, and students' psychological learning processes when studying the effectiveness of a TML system (i.e., when studying the ability of TML to generate effective learning outcomes).

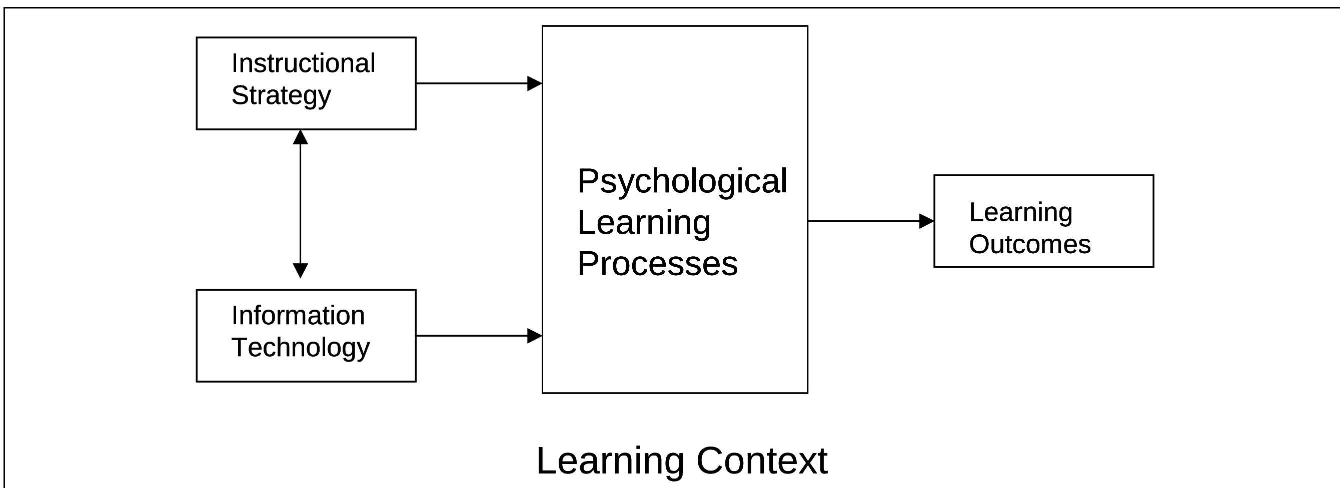

Figure 1. A Framework for TML Research
(Reprinted by permission, M. Alavi and D. E. Leidner, "Research Commentary: Technology-Mediated Learning—A Call for Greater Depth and Breadth of Research," *Information Systems Research* (12:1), 2001, p. 5. Copyright © 2001, The Institute for Operations Research and Management Sciences, 7420 Parkway Drive, Suite 310, Hanover, Maryland 21076.)

Learning Theories and Instructional Strategies

As the framework of TML research indicates, instructional strategies can significantly affect the effectiveness of information technology in learning. The instructional strategies, on the other hand, are mostly drawn from learning theories. Among the many existing learning theories, there are three theories that especially emphasize the importance of interaction in learning: constructivist theory (Bruner 1966), discovery learning theory (Bruner 1960), and conversation theory (Pask 1975).

The first of these three theories, the constructivist theory, is broadly applied in both traditional classroom and computer-supported learning. Constructivist theory asserts that learning is an active process in which learners construct new ideas or concepts based upon their current and past knowledge (Bruner 1966). This theory emphasizes that learning should be "learner-centered" instead of "teacher-centered" as in traditional classroom training. The instructor and students should engage in an active dialog, while guidance and direct instruction should only be provided when necessary (Phillips 1995).

Similarly, according to discovery learning theory, students should be provided with data and be called upon to question, explore, and/or experiment to discover knowledge. For example, students may be expected to discover the particular principle(s) contained in a lesson by questioning the teacher. As a result of the discovery process, it is premised that students will be better able to remember and apply what they have learned (Bruner 1960).

Finally, the fundamental idea of conversation theory (Pask 1975) is that learning occurs through conversations about a subject. These conversations serve to make the learner's knowledge explicit. In the human learning process, natural language conversations such as questioning and answering are the major type of conversation, facilitating learning through exploration and reflection (Schank and Cleary 1995).

These three modern learning theories all indicate that high quality learning should incorporate active interactions initiated by students, such as the questioning and answering process. This suggests that the instructional strategies implemented in an effective e-learning system should facilitate active interactions in learning.

However, Jonassen (1991) states that each phase of knowledge acquisition requires different types of learning. Initial knowledge acquisition is perhaps best served by classical instruction, while a constructivist learning environment is more suited to the second phase of knowledge acquisition in which learners acquire more specific, advanced knowledge to solve more complex, domain-specific problems. Therefore, as the literature on constructivist instructional design suggests, a mix of instructional design strategies, combining both learning from traditional classroom instruction and conversation-based constructivist learning, may benefit learners and meet the needs of a variety of learning situations better than any one strategy (Ertmer and Newby 1993). For example, question-based learning can be used for just-in-time learning after initial, traditional training on the subject matter is completed.

In addition to emphasizing active dialog for constructing new knowledge, constructivist theory also emphasizes the importance of past knowledge and feedback. Appropriate assessment of a learner's prior knowledge and the respective feedback after the assessment guide the learner in deciding how to interpret the realities, and whether to confirm, refine, or revise their mental models (Guskey 1997). Broadly speaking, interaction in learning should also include such an assessment and feedback process.

Currently, interactions in the TML environment can be classified into three types (Moore 1993): (1) learner–content interaction (LCI), (2) learner(s)–tutor(s) interaction (LTI), and (3) learner(s)–learner(s) interaction (LLI). LCI refers to students questioning the training materials. Of the three types of interaction, LCI has been the least studied, perhaps because it is often assumed that one can only interact with a person. However, Bates (1995) states that interactions occur not only with the originator of knowledge, but also with knowledge itself, as represented in different media. Furthermore, many research studies have shown that thoughts can be triggered, meaning can be made, and mental models can be created, revised, and recreated for learners through their interacting with information received and revisited (Anderson and Pearson 1984; Piaget 1967; Schallert 1982). In contrast, both LTI and LLI are interactions among people. These types of interactions have the advantages of increasing the involvement of students in learning, as well as bringing more personal and social aspects to the interactions, which cannot be supported by simple LCI (Hiltz and Wellman 1997). Existing research studies focus primarily on LTI and LLI, especially in the research of computer-supported collaborative learning. However, one disadvantage of LTI is its reliance on the availability of live tutors, because the resource of qualifying tutors may be scarce and/or mentors are often not available at the exact time needed. Although asynchronous interaction may provide a temporary solution when the tutor or instructor is unavailable, asynchronous interaction cannot always provide the timely feedback that is generally desirable to support learning.

Because of the disadvantages of both LCI and LTI, this research focuses on investigating technologies to support a special type of LCI, called *virtual interaction*. Virtual interaction simulates LTI with a virtual mentor or tutor, thus it does not rely on the availability of live mentors or tutors. The next section describes technology that strives to simulate LTI and to support virtual interaction.

Core Technology for Virtual Interaction

Simulating LTIs without reliance on a human tutor was difficult to implement just a few years ago. Thanks to a new technology, video-based question answering (QA), virtual interaction is now possible. By prerecording a mentor's instructions (e.g., lectures) in digital video format, the human questioning and answering process can be simulated by a process of finding the specific video segment that is most relevant to the student's question. Because students can ask specific questions in natural language and watch and hear a person responding, such virtual interactions simulate LTI.

The video-based QA technology integrates speech recognition, information retrieval, and natural language processing technologies. Zhang and Nunamaker (2004) proposed a method that applies text-based QA approaches to video applications. In this approach, the speech in video lectures is transcribed into text manually or automatically using speech recognition software. Each long video of a lecture is also manually segmented into short segments representing lecture topics. The transcript of each video segment is treated as a text document. A template-based approach is then used to identify answers to posted questions from the collection of transcribed video segments. After (1) parsing the posted question into major verbs, nouns, noun phrases, and named entities using a parser called Connexor iSkim (Voutilainen 2000), (2) obtaining their synonyms from the WordNet dictionary (Miller 1990), and (3) deriving the answer type of the question (e.g., "Who + BE," "Where + BE," etc.) based on a set of linguistic rules, this template-based approach fills the posted question into a question template with nine slots (Zhang and Nunamaker 2004). The video transcripts are parsed in the same fashion to generate sentence templates, and the question template is compared with sentence templates in the video transcripts to retrieve the most relevant video segments.

The described template-based approach was used as the core technology to implement virtual interaction, combined with phonetic matching technology and additional knowledge sources to improve the performance of the virtual interactions. The proposed new approach is described in detail in Cao et al. (2005).

Finally, although video segmentation technique and its facilitation in e-learning exist in many commercial-off-the-shelf knowledge management systems, the natural language and video-based virtual interaction technology is still in its very early stage. Therefore, the impacts of this new technology on learning are still to be explored and are the focus of the research described in this paper.

The LVM Model and Prototype

To investigate the impacts of virtual interaction on learning, we propose an exploratory, conceptual model, Learning with Virtual Mentors (LVM, Figure 2). Based on the Alavi and Leidner (2001) framework of TML research, we propose the LVM model with a comprehensive view of the relationships among instructional strategies, information technologies, students' psychological characteristics, environmental factors, and learning outcomes.

In the LVM model, the learning outcomes are represented by learners' actual learning performance and their perception of and satisfaction with the whole learning process (e.g., perceived learning effectiveness and self-reported interactivity). The learning outcomes may be affected by several factors related to different instructional strategies, information technologies, and students' psychological characteristics. For example, learning phases, students' learning motivations and strategies, learner capabilities, system interactivity, and learning activities are all possible factors influencing learning outcomes. Although a series of studies have been conducted to explore the relationships among all constructs in the LVM model depicted in Figure 2, because of space limitation this paper only focuses on describing part of the studies about the following core constructs.

Learning Phases. As stated earlier, different learning phases require different instructional strategies. Matching the appropriate instructional strategy to the learning stage may greatly impact the learning outcomes (Jonassen et al. 1993). The first phase, introductory learning, occurs when learners have little or no prior knowledge about a content area, while in the second phase of learning, learners acquire more specific, advanced knowledge to solve more complex, domain-specific problems. Similar to what Ertmer and Newby (1993) stated, we propose that the introductory learning phase is best served by classical instruction, while interactive learning is more suitable to the advanced learning phase. In particular, we believe that questioning and answering learning strategies will support the advanced learning phase.

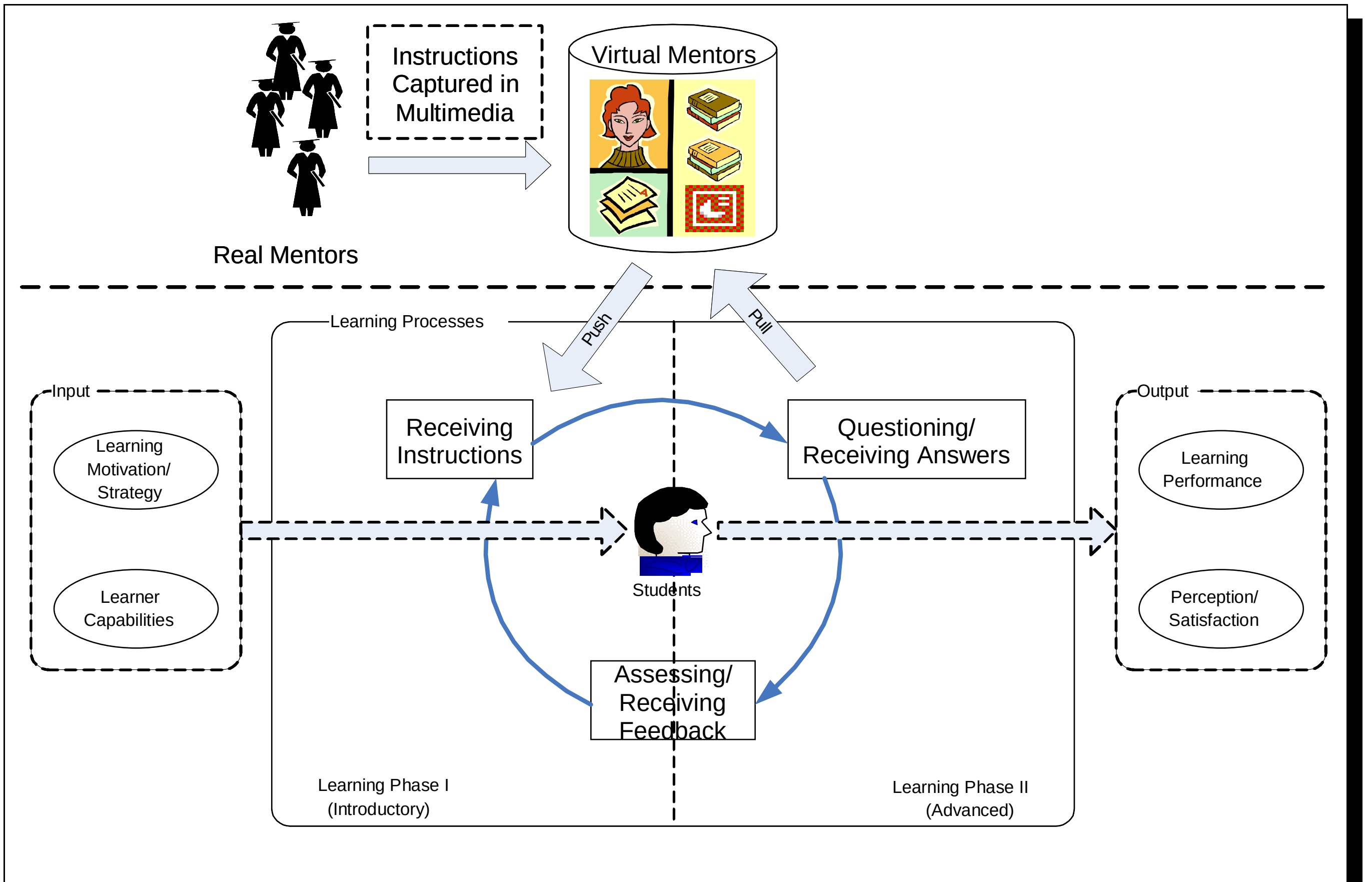

Figure 2. The LVM Model

System Interactivity. Affected by both the instructional strategy and information technology, system interactivity is the key construct in the LVM model. As shown in Figure 2, the classical, sequential instruction provides zero interaction in learning; the question and answering process initiated by students adds interactivity into the student learning process; and the assessment and feedback process initiated by the LVM further increases the interactivity. Therefore, it is expected that more virtual interactive functions available in an e-learning system will correlate to more students' interactive learning activities, which will result in improved student learning outcomes. However, system interactivity may not play a big role in the introductory learning phase when students focus on gathering initial knowledge through traditional instruction. We expect that the effect of system interactivity will show up in the advanced learning phase when students are trying to deepen their understanding of the subject matter.

Learning Activities. The learning activities construct refers to learners' interactive activities in learning. Although it is difficult to measure learners' actual interactive activities in the traditional classroom learning environment, it is easy to record them in system logs in an e-learning system. We propose that more interactive learning activities can be the result of more system interactivities, resulting in better learning outcomes in the advanced learning phase.

Although individual learner characteristics such as learning motivation and learner capability are also important components of the LVM model, they are not investigated herein. Instead, they are controlled by random assignment in the experimental studies. The focus of this research is testing the following four hypotheses about the core constructs of learning phases, system interactivity, and learning activities.

Hypothesis 1: Higher levels of system interactivity (assuming a real mentor has the highest level of interactivity) result in students engaging in more interactive learning activities in the advanced learning phase, as well as in the overall learning process.

Hypothesis 2: Students will engage in more interactive learning activities in the advanced learning phase than in the introductory learning phase.

Hypothesis 3: Higher levels of system interactivity result in better learning outcomes (A. Actual learning performance; B. Perception or satisfaction) in the advanced learning phase, as well as the overall learning process.

Hypothesis 4: Increasing the level of system interactivity in the advanced learning phase results in a greater improvement in learning outcomes (A. Actual learning performance; B. Perception or satisfaction) than in the introductory learning phase.

To test the above hypotheses about the LVM model, a Web-based prototype was developed using advanced information technologies, including the core video-based QA technology. As illustrated in Figure 3, in the LVM prototype with the most interactive functions, the Web-based user interface is divided into four cells or sections: (1) a video display of the instructor, (2) a PowerPoint slide associated with the current video segment, (3) a text note, which is the transcription of the speech in the current video segment, and (4) an outline of all the topics in the lecture. The content in the four cells is synchronized. Furthermore, students may interact with the LVM system in two ways. First, each topic in the outline is directly linked to the relevant video segment, allowing students to click on any link to review a specific topic. The outline can be viewed as a pre-compiled list of questions, although it is less flexible than the second way of virtual interaction: direct QA. A textbox in the center of the four cells is provided for students to ask questions. Once a question is submitted, a new window with the four-cell design will pop up to present the potential answers. The video segment determined to be the "best answer" will be automatically played, with its associated PowerPoint slide and text note appearing in the other cells. Therefore, students can immediately watch and hear the virtual mentor responding to their questions. In the answer window, the lecture outline is replaced with the list of answer topics. If students are not satisfied with the first answer, they can click on the links in this answer list to view the other video answers.

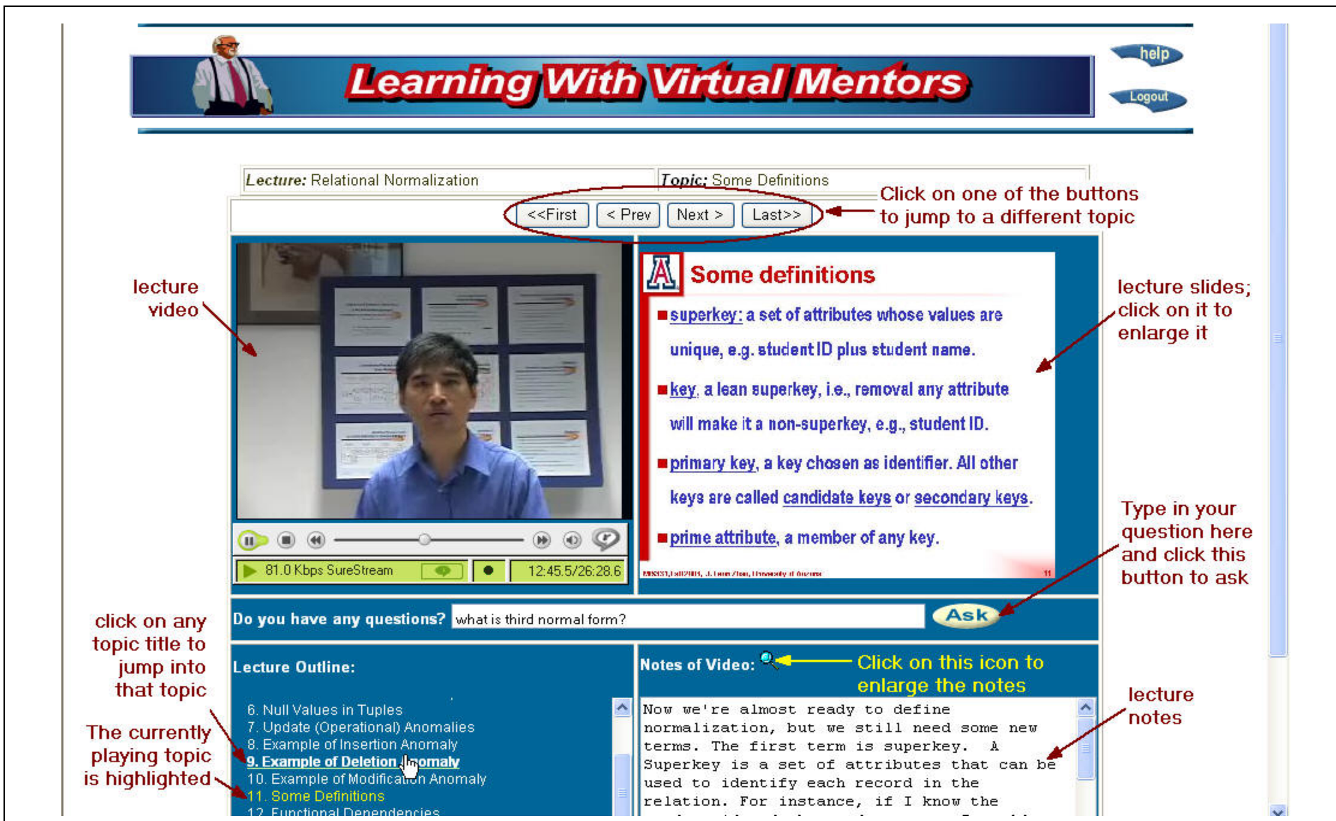

Figure 3. User Interface of the Prototype System

Besides the interactions initiated by students, the virtual mentor uses pop-up quizzes to initiate interaction with students. Pop-up quizzes are implemented in this LVM prototype and a QA box is added into the quiz feedback page, so that if students find their answer is not correct, they can immediately ask the virtual mentor a question.

Several video lectures have been created to be used in this prototype. In the study described in this paper, a lecture (about 40 minutes long) about computer security concepts was used.

Study Design

When investigating the LVM model, we conducted a field experiment in conjunction with observations such as survey and interviews. To explore the relationship among the learning outcomes and the factors such as system interactivities and learning phases, an experiment was conducted as a longitudinal pretest-posttest comparison between control and treatment groups (Table 1). Since the relationship between the system interactivities and the learning outcomes was our major interest, we arranged five treatment groups in which participants used the LVM prototype with different types and numbers of interactive functions (outline, QA and/or quizzes), and one control group in which a real mentor taught the participants face-to-face.

Participants

Participants (N = 158) were undergraduate students (59 percent male, 41 percent female) recruited from an Information Systems course at a southwestern university. They volunteered to participate in this research study as an alternative of an assignment (the other alternative was to finish a research report for a class-related topic). They were told that the learning materials presented in this study were more advanced knowledge of the topics (computer security) they were studying in the IS course. Students could get full credits for this assignment as long as they actively participated in all experimentation sessions. Of the students participating, 69 percent reported their grade point average (GPA) as between 3.0 and 4.0 on a 4.0 scale, while 31 percent of the students reported their GPA as between 2.0 and 3.0.

Procedures

Participants signed up for this research study through a Web site and chose to attend one of six available time slots. The six time slots were randomly assigned to the six groups (five treatment groups and one control group). All groups used the same computer lab for the research study, including the Real Mentor group (the control group).

Table 1. Experimental Design

<table>
<tr><th rowspan="2"></th><th colspan="6">Treatments</th></tr>
<tr><th>Real Mentor</th><th>VM – IN only</th><th>VM – IN+ Outline</th><th>VM – IN+QA</th><th>VM – IN+QA +Outline</th><th>VM – IN+Outline +QA +Quiz</th></tr>
<tr><td rowspan="3">Class session</td><td></td><td colspan="5">Pretests (45 minutes)</td></tr>
<tr><td colspan="6">Learning (50 minutes, including 10 minutes of review time)</td></tr>
<tr><td></td><td colspan="5">Posttests (30 minutes)</td></tr>
<tr><td colspan="7">Two Weeks</td></tr>
<tr><td rowspan="3">Assignment session</td><td></td><td colspan="5">Delayed posttest (15 minutes)</td></tr>
<tr><td></td><td colspan="5">Finishing an assignment (50 minutes)</td></tr>
<tr><td></td><td colspan="5">Posttests (30 minutes)</td></tr>
</table>

The study began with a researcher (one of the authors) describing the study design and procedures to the student participants. Students in each of the five treatment groups were then given a quick demo of the particular system (treatment) they would use. Students in each group participated in two learning sessions with the same system (treatment). They were not allowed to switch to a different treatment group in the second session. The first session was designed to be introductory, non-task-oriented learning. In this session, participants first filled out a preexperiment survey (30 minutes total, including demographic information and survey questions measuring students' learning motivation, strategy, and styles), and then completed a pretest that had 10 multiple-choice questions about the computer security concepts (15 minutes). Students then received the computer security lecture in their particular treatment conditions (50 min; with human instructor or using the LVM prototype). Finally, they completed a posttest, which had the same questions as those in the pretest, but the order of both the questions and the answer choices were reordered (15 minutes). After the posttest, the participants answered a post-experiment questionnaire (15 minutes, including questions about perceived learning effectiveness and other factors). In addition, the participants' activities when using the LVM system were recorded to a system log file, and the activities of the participants in the "real mentor" group were videotaped for future analysis.

The second session was designed to be advanced, task-oriented learning. Two weeks after the first session, the same participants were asked to attend the second session and they were required to learn with specific tasks. Specifically, they were required to complete an assignment with more in-depth questions about the computer security concepts. They first took a 15-minute delayed posttest on their knowledge about computer security, then they completed an open-ended assignment in 50 minutes with the help of either asking the instructor directly (in the control group; only one student can ask a question at a time) or using the system (in the other five treatment groups), and then they completed another posttest (15 minutes). Finally, participants completed the same post-experiment questionnaire as in the first session (15 minutes). Again, during the 50-minute assignment session, the participants' activities in using the LVM training system were recorded to a system log file, and the activities of the participants in the "real mentor" group were videotaped for future analysis.

Measures and Instruments

Independent Variables for the Experiment

System Interactivities. Participants were randomly assigned to one of the six groups in which they would have access to different levels of interactivities (see Table 1). We assumed that a real mentor could provide the highest level interactivity in the learning process. The five treatment groups are listed below with a brief description of the system interactivity for each. The system interactivity increases from treatment group 1 to group 5.

Treatment 1. VM - IN only: students could only watch the instruction sequentially.

Treatment 2. VM – IN + Outline: students could click on the links in the outline to change topic when watching the instruction.

Treatment 3. VM – IN + QA: students could ask questions when watching the instruction.

Treatment 4. VM – IN + Outline + QA: students could not only click on the links in the outline to change topic but also ask questions when watching the instruction.

Treatment 5. VM – IN + Outline + QA + Quiz: students could not only click on the links in the outline to change topic but also ask questions when watching the instruction; In addition, the system could pop-up questions in the middle of the instruction.

Learning Phases. Each participant would go through two learning phases. The first one, class session, is the introductory learning phase. The second one, assignment session, is the advanced learning phase.

Dependent Variables

Learning Activities. This variable was measured for the five treatment groups where all students' learning activities were recorded into a system log file. There were three types of interactive activities recorded in the system: (1) asking a question, (2) switching topic using the outline, and (3) switching topic using the navigation buttons. This variable was measured by the total number of the student's interactive activities during a session.

Learning Performance. The student's actual learning performance was measured by his or her percentage accuracy score on the posttests (10 multiple-choice questions on the student's knowledge of computer security), as compared to the score of the pretest. Students took the same format of tests four times in 2 sessions. The order of the questions and the response choices were alternated for each test.

Perceived Learning Effectiveness. The student's perceived learning effectiveness **(EFFECT)** was measured by a scale consisting of eight items in the post-experiment questionnaire adapted from Alavi (1994) ($\alpha = .88$ in session 1 and $\alpha = .91$ in session 2). For all items in the post-experiment questionnaire, students rated themselves on a five-point Likert scale ranging from "strongly disagree" (scale = 1) to "strongly agree" (scale = 5). An individual's score on the scale was conducted by taking the mean of the eight items that made up that scale. A high score on this scale meant that the student thought he or she learned effectively in this study.

Self-reported Interactivity. This variable **(INTER)** concerned students' satisfaction with their interaction with the virtual mentor (or real mentor in the control group). It was measured by a self-developed, four-item scale in the post-experiment questionnaire ($\alpha = .82$ in session 1 and $\alpha = .88$ in session 2). Again, an individual's score on the scale was conducted by taking the mean of the items that made up that scale. A high score on this scale meant that the student felt that he or she had interactions with the (virtual) mentor and enjoyed the (virtual) interactions.

Analysis and Results

Besides the quantitative data collected in the experiment, qualitative data was collected from open-ended questions in the post-experiment questionnaires and was used to help interpret the quantitative results. However, in this paper we focus on discussing the quantitative results and only use the general findings from the qualitative analysis to explain the results.

Results about Interactive Learning Activities

Table 2 lists the means and standard deviations of students' interactive learning activities in each session.

As mentioned earlier, all students' learning activities were recorded into a system log file for the five treatment groups. However, in the first group, the "IN only" group, the students had no access to the interactive system functions such as QA and outline. Therefore, only four groups had interactive learning activities recorded in system logs and our analysis was based on these four groups.

Repeated measures analysis indicated that students' interactive learning activities were different for the factor *Session*, $F(1, 103) = 204.776$, $p < .005$, partial $\eta^2 = .665$; and interaction *Session*Treatment*, $F(3, 103) = 4.097$, $p = .009$, partial $\eta^2 = .107$. The *Session*Treatment* interaction indicated that although all groups had more interactive learning activities in the advanced learning session (hypothesis 2 was supported), the increase was different among groups. A reverse Helmert contrasts analysis was conducted on the *Treatment* factor. Results indicated that students in group 3 engaged in significantly more interactive activities

Table 2. Means and Standard Deviations of Interactive Learning Activities

		Session 1 Activities		Session 2 Activities
Group	N	mean (std)	N	mean (std)
2 (IN + Outline)	25	6.44 (6.01)	25	19.28 (10.50)
3 (IN + QA)	24	2.92 (2.90)	23	27.54 (8.71)
4 (IN + QA + Outline)	28	9.04 (4.95)	28	25.04 (14.14)
5 (IN + QA + Outline + quizzes)	30	6.50 (7.58)	30	22.20 (13.41)
Total	107	6.35 (6.05)	106	23.46 (12.27)

than students in group 2 in session 2 (p = .018); however, the greater number of system interactivities in groups 4 and 5 did not significantly increase the actual interactive learning activities in this advanced session. A *post hoc* pair-wise comparison of the interactive activities in session 2 did not show significant differences among groups 3, 4,and 5. Therefore,hypothesis 1 was only partially supported. After carefully examining the study procedure, we found that the students' activities of answering the pop-up quizzes (the only different functionality between groups4 and 5) were accidentally not recorded in the system log, even though these activities were mandatory (i.e., the students were required to answer the pop-up questions before they could continue). Therefore, it appears reasonable that groups 4 and 5 had similar recorded activity patterns. Since the major difference between group 2 and groups 3, 4, and 5 was the QA function, this may indicate that it was the QA function that triggered more interactive learning activities in the advanced learning session; however, more research needs to be conducted for confirmation of this result.

Another important observation was that, although we assumed that the real mentor would be able to provide the most amount of interactivities, in both of the sessions we found that there were very few interactions between the students and the instructor in the classroom. In the first session, two students asked the live instructor a total of three questions at the end of the lecture, while in the second session, only four students asked the instructor questions when completing the assignment. This was obviously different from the expectation. Our qualitative results from the semi-structured interviews after the experiment helped us explain this circumstance. When asked, "Why didn't you ask the instructor questions when you did not understand the topic?," the most common responses were "I think maybe another student would ask this question," "I don't want to appear stupid if the other students all understand what I wanted to ask," "I don't want to disturb the whole class," and "I don't want to wait in the line to ask a question." Therefore, combined with our findings about the interactive activities in the treatment groups, we found that virtual interactions could actually remove this psychological barrier and enable more direct question and answering processes in learning.

Results about Learning Performance

Table 3 lists the means and standard deviations of students' learning performance test scores in each session. Although percentage scores were presented here for easier interpretation, ArcSin transformation was performed on all percentage scores to improve the equality of variance before conducting any ANOVA or ANCOVA test.

A 6 × 2 × 2 repeated measures analysis was conducted to test the hypotheses. Results showed that students' learning performance test scores were significantly different for the factor *Session*, $F(1, 146) = 126.467$, $p < .005$, partial $\eta^2 = .464$; and interaction *Session*Prepost*, $F(1, 146) = 20.706$, $p < .005$, partial $\eta^2 = .124$. This indicated that students in all groups had higher learning performance test scores in the advanced learning session, but the test score gains were different from session 1 to session 2. All groups had more test score gains in the first session than in the second session. Hypothesis 4A was not supported. This is quite understandable because in session 2 there was less space for improvement. However, no significant treatment effect was found through the repeated measures analysis. It seemed that all groups had similar patterns in test score changes. Hypothesis 3A was not significantly support either.

Table 3. Means and Standard Deviation of Learning Performance

	Session 1 (Class Session)			Session 2 (Assignment Session)		
		Pretest	Posttest		Pretest	Posttest
	N	mean (std)	mean (std)	N	mean (std)	mean (std)
IN only	26	28.8 (16.1)	72.7 (16.4)	22	50.9 (16.9)	80.5 (16.5)
IN + Outline	25	29.2 (17.1)	68.8 (17.2)	25	43.2 (17.3)	74.0 (19.8)
IN + QA	24	26.3 (15.0)	73.3 (17.4)	23	48.3 (19.7)	80.4 (14.3)
IN + QA + Outline	28	32.5 (21.0)	71.8 (17.7)	28	51.4 (18.6)	76.4 (14.5)
IN + QA + Outline + quizzes	30	29.0 (14.7)	66.3 (19.2)	30	44.7 (21.3)	76.7 (18.8)
Real Mentor	25	30.4 (21.7)	68.0 (18.3)	24	45.4 (21.3)	65.8 (20.2)
Total	158	29.4 (17.6)	69.9 (17.6)	152	47.2 (19.3)	75.6 (17.9)

As suggested by Bonate (2000), a more powerful ANCOVA test was then conducted to see if there was really no difference among treatment groups. Since we hypothesized that the performance gain would be different in the second session and in overall learning, we conducted two ANCOVA tests respectively. One took the posttest in session 2 as dependent variable and the other took the posttest in session 1 as dependent variable. Both tests took pretest in session 1 as covariate.

The results were indeed different from the results of the repeated measures analysis. There were significant differences among groups in both session 2 ($F(5, 145) = 2.274$, $p = .05$, partial $\eta^2 = .073$) and the overall process ($F(5, 145) = 2.409$, $p = .039$, partial $\eta^2 = .077$). In session 2, a reverse Helmert contrasts analysis revealed that the real mentor group had significantly less test score gain in comparison to the mean of the five system groups. Similar results were found for the overall learning process. The overall test score gain (posttest in session 2, pretest in session 1) in the real mentor group was significantly lower than the mean of the other five system groups.

Therefore, the real mentor group, which was expected to have the highest learning performance gain in both session 2 and overall (see hypothesis 3A), actually had the lowest learning performance gain. Based on our observations about the students' actual interactive learning activities in both sessions, we might argue that the real mentor in a classroom setting would have the least interactivities and thus hypothesis 3A might still be partially supported. However, more studies are needed for retesting this hypothesis.

Results about Perception and Satisfaction

Table 4 lists the means and standard deviations of the two variables from the self-report post-experiment survey in each session.

A multivariate ANOVA test on the EFFECT and INTER measures showed that students' perceived learning effectiveness and perceived interaction with the (virtual) mentor had weakly significant difference between the two sessions, Wilk's $\lambda = .961$, $F(2, 143) = 2.938$, $p = .056$, $\eta2 = .039$. A follow up univariate analyses about students' perceived interactivity found no significant difference between sessions but significant difference among treatment groups (between-subject main effect, $F(5,144) = 12.490$, $p < .005$). A reverse Helmert contrast analysis found that the perceived interactivity was significantly higher in the real mentor group than in the other five treatment groups both in session 2 ($p < .005$) and in the overall learning process ($p < .005$). It also found that treatment group 3 (IN+QA) had significantly higher perceived interactivities than the mean of group 1 and group 2, both in session 2 ($p = .001$) and in the overall learning process ($p = .010$). However, there was no significant difference among groups 3, 4, and 5. Therefore, hypothesis 3B was only partially supported. This was outside our expectations, but was quite similar to what happened with the actual interactive learning activities.

Therefore, the students' perceived (self-reported) interactivity is quite different from their actual interactive learning activities. First, students' perceptions about interaction did not change along with their actual behavior across sessions. Second, although many students did not talk with the instructor at all in the classroom, they still thought they had good interaction with the instructor; on the other hand, while some students did ask more questions in the LVM system, they did not feel they were inter-racting. The qualitative results by analyzing the open-ended responses confirmed this finding. On one side, students' under-

Table 4. Means and Standard Deviations of Self-Reported Learning Outcomes

	Session 1 (Class Session)			Session 2 (Assignment Session)		
		EFFECT	INTER		EFFECT	INTER
	N	mean (std)	mean (std)	N	mean (std)	mean (std)
IN only	26	3.64 (.50)	2.68 (.82)	22	3.57 (.66)	2.69 (.75)
IN + Outline	25	3.56 (.53)	2.91 (.77)	25	3.24 (.96)	2.43 (.92)
IN + QA	24	3.68 (.68)	3.04 (.76)	22	3.54 (.58)	3.23 (.63)
IN + QA + Outline	28	3.47 (.68)	2.94 (.84)	28	3.55 (.69)	3.02 (.84)
IN + QA + Outline + quizzes	30	3.65 (.58)	3.06 (.66)	29	3.34 (.95)	2.83 (.90)
Real Mentor	25	3.86 (.72)	3.98 (.55)	24	3.68 (.69)	3.98 (.50)
Total	158	3.64 (.62)	3.10 (.83)	150	3.48 (.78)	3.02 (.91)

standing of interaction in learning was a bit different from what we expected. They would count the potential or capability of interaction with a live instructor to be good interaction. Many students implied that interaction, in their view, had to be between live people. The nonverbal communication features such as eye contact or gestures, as well as the feeling of social presence, might contribute to this circumstance. On the other side, many students did not perceive the virtual interaction as good interaction because of a limitation of the current technology: the answers were repeated parts of the prerecorded lecture. Therefore, most students quickly realized that the virtual mentor could not rephrase the answer based on their individual need, nor could it answer anything out of the boundary of the lecture. Because of this key difference between virtual interaction and the interaction with a live instructor, some students did not accept virtual interaction as real, two-way interaction.

The qualitative results also confirmed that virtual interaction with QA is better than interaction with outline (groups 3, 4, and 5 had more actual interactive activities and perceived interactivity than group 2), especially in session 2. Students' satisfaction with outline decreased from session 1 to session 2, while their satisfaction with QA increased from session 1 to session 2. Also, in session 2, more students thought QA was better than outline in finding specific information. Therefore, it seems that outline is more suitable for the introductory learning session as it provides a good map for students building their own mental models, while QA is more important for the advanced learning process when students need reinforcement on specific topics.

A univariate analyses about students' perceived learning effectiveness, on the other hand, revealed that it was significantly less in session 2 than in session 1, $F(1,144) = 5.914$, $p = 0.016$. Therefore, hypothesis 4B was not supported. One possible explanation for this finding is that students judged their learning effectiveness only on the amount of new knowledge they received. Because students learned the same content twice, they did not feel they learned new knowledge in the second session and, therefore, had less perceived learning effectiveness. A reverse Helmert contrast analysis found that overall the perceived learning effectiveness was significantly higher in the real mentor group than in the other five treatment groups ($p < .005$). However, we failed to find significant difference among the five treatment groups both in session 2 and in the overall learning process. Hypothesis 3B was only partially supported.

Again, similar to the findings about interactivity, the students' perceptions are the opposite of the reality. Although students in the real mentor group had the least actual interactive learning activities and had the lowest learning performance gain, they had the highest perceived learning effectiveness. We also failed to find significant difference among the five treatment groups for both the actual and perceived learning effectiveness. Looking at the results from qualitative observations, we realized that the self-paced control and convenience of reviewing specific content commonly provided in all system groups are actually the key factors that contribute to the difference between the effectiveness of the virtual mentor and the real mentor in the advanced learning session. Different levels of interactivity actually determine different levels of convenience of reviewing specific content; thus, as long as this convenience exists, students may learn better with the virtual mentor than with a real mentor. The benefits are more from the satisfaction perspective. For example, when QA functions were available, students included QA as one of the reasons they liked the system, while only a few students included the outline or the pop-up quiz as one of the reasons. In addition, two or three times more students in the second session listed QA as one of the reasons they liked the system better than in the first session. This again confirmed that interaction, especially QA, is more important for the advanced learning phase.

Summary and Discussion

To summarize, this research studies how information technologies, such as automatic question answering (QA), can be used to provide virtual interaction in a multimedia-based e-learning system and turn it into a virtual mentor to provide students with interactive, one-on-one instruction. More importantly, this study explores the key factors that may affect the effectiveness of a virtual mentor, which are proposed in the exploratory LVM model. Specifically, this paper focuses on describing part of a series of studies that explore the relationships among some of the core constructs in LVM.

Results from experiments and observations were summarized in Table 5, which shows that, as expected, learning phases do impact students' learning behaviors to some extent; all groups had more interactive learning activities in the advanced learning session. However, students' perceived (self-reported) interactivity is very different from their actual interactive learning activities. In particular, most students still perceived more interactivities with a human mentor than with a virtual mentor system, even if they actually interacted more with the virtual mentor. The possible reasons for this discrepancy include both the limitation of the current technology and the social aspects of interaction. We think that the technology limitations might be mitigated by extending the answers with content extracted from the Web; however, the social aspect of interaction with a live instructor will be difficult to fully simulate by virtual interaction with current technology. Asynchronous collaboration may still be needed to complement the virtual interaction.

<table>
<caption>Table 5. Summary of Results</caption>
<tr><th colspan="3">Hypotheses</th><th>Results</th></tr>
<tr><td rowspan="2">Activity</td><td colspan="2">H1 (higher levels of system interactivity result in students engaging in more interactive learning activities)</td><td>Partially Supported. QA group engaged in more interactive activities than the Outline group.</td></tr>
<tr><td colspan="2">H2 (more interactive learning activities in the advanced learning phase than in the introductory learning phase)</td><td>Supported.</td></tr>
<tr><td rowspan="4">Outcome</td><td rowspan="2">H3 (higher levels of system interactivity result in better learning outcomes)</td><td>A. Learning Performance</td><td>Not supported. Real mentor group had the lowest test score gain. No significant differences among other groups.</td></tr>
<tr><td>B. Perception/ Satisfaction</td><td>Partially supported. Real mentor group perceived the most learning effectiveness and interaction; QA+Outline group perceived more interaction than the QA or Outline groups.</td></tr>
<tr><td rowspan="2">H4 (more interactivity in the advanced learning phase results in a greater improvement in learning outcomes than in the introductory phase)</td><td>A. Learning Performance</td><td>Not supported. Less test score gain in the advanced learning phase; and the change of learning gains between two phases is not significantly different among treatment groups.</td></tr>
<tr><td>B. Perception/ Satisfaction</td><td>Not supported. Less perceived learning effectiveness and perceived interaction in the advanced learning phase; and the change of learning gains between two phases is not significantly different among treatment groups.</td></tr>
</table>

Although students in the virtual mentor groups did not perceive as much interaction as those in the real-mentor group, students who were provided with the QA function did have significantly higher perceived interactivities than those who were only provided with the outline function. Overall, the QA type of virtual interaction did enable more actual interaction and perceived interactivity than the traditional hyperlink type of interaction when students needed more interactions in the advanced learning session.

Although the virtual interaction technology did impact students' behaviors in interaction and improve students' satisfaction with e-learning to some extent (especially in the advanced learning phase), its impact on learning effectiveness is limited. The factors "pace control" and "convenience for access information" may also be key factors that directly affect the learning effectiveness of an e-learning system with virtual interactions.

The research findings reported in this paper have both research and practical implications. The research studies the effectiveness of e-learning with a focus on the impact of interaction, particularly a new type of interaction defined as virtual interaction, which is still learner-content interaction (LCI) but simulates the learner-tutor interaction (LTI). Since virtual interaction does not rely on the availability of human instructors, it can be more cost efficient than using human instructors. Although virtual interaction was not found to directly impact learning performance, it is somewhat related to students' satisfaction with e-learning. We hope the addition of virtual interaction will be able to reduce the drop rate of e-learning and attract more students.

Finally, findings from this research provide suggestions for e-learning practitioners. The technology features of e-learning need to be adjusted based on different learning phases in order to achieve the best learning outcomes. Specifically, in the introductory course, using simple hypermedia lectures is fine and can reduce cost and instructor workload, but in the advanced learning phase, which reviews and reinforces the old content, the virtual interaction should be added to help students quickly find answers to their questions. Based on the observations in this study, virtual interaction may not be sufficient for a more advanced, discussion type of class, and asynchronous collaboration or face-to-face classroom discussion may still be needed. In this situation, an appropriate combination of virtual and real mentor(s) may prove to be the best way to learn.

References

Alavi, M. "Computer-Mediated Collaborative Learning: An Empirical Evaluation," *MIS Quarterly* (18:2), 1994, pp 159-174.

Alavi, M., and Leidner, D. E. "Research Commentary: Technology-Mediated Learning—A Call for Greater Depth and Breadth of Research," *Information Systems Research* (12:1) 2001, pp. 1-10.

Anderson, R. C., and Pearson, P. D. "A Schema-Theoretic View of basic Processes in Reading Comprehension," in *Handbook of Reading Research*, P. D. Pearson (Ed.), Longman, New York, 1984.

Bates, A. W. *Technology, Open Learning and Distance Education*, Routledge, London, 1995.

Bonate, P. L. *Analysis of Pretest-Posttest Designs*, Chapman & Hall/CRC, London, 2000.

Bruner, J. *The Process of Education*, Harvard University Press, Cambridge, MA, 1960.

Bruner, J. *Toward a Theory of Instruction*, Harvard University Press, Cambridge, MA, 1966.

Cao, J., Roussinov, D., Robles-Flores, J. A., and Nunamaker, J. F. "Automated Question Answering From Videos: NLP vs. Pattern Matching," in *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, IEEE Computer Society Press, Los Alamitos, CA, January 3-6, 2005, pp. 43-53.

Ertmer, P. A., and Newby, T. J. "Behaviorism, Cognitivism, Constructivism: Comparing Critical Features from an Instructional Design Perspective," *Performance Improvement Quarterly* (6:4) 1993, pp. 50-70.

Guskey, T. R. *Implementing Mastery Learning* (2nd ed.), Wadsworth Publishing Company, Belmont, CA, 1997.

Hiltz, S. R., and Wellman, B. "Asynchronous Learning Networks as a Virtual Classroom," *Communications of ACM* (40:9), 1997, pp. 44-49.

Jonassen, D. H. "Objectivism Versus Constructivism: Do We Need a New Philosophical Paradigm?," *Educational Technology Research and Development* (39:3), 1991, pp. 5-14.

Jonassen, D. H., McAleese, R., and Duffy, T. M. "A Manifesto for a Constructivist Approach to Technology in Higher Education," in *The Design of Constructivistic Learning Environments: Implications for Instructional Design and the Use of Technology*, T. M. Duffy, J. Lowyck, and D. H. Jonassen (Eds.), Springer-Verlag, Heidelburg, Germany, 1993.

Miller, G. "WordNet: An On-line Lexical Database," *International Journal of Lexicography* (3:4), 1990, pp. 235-244.

Moore, M. G. "Three Types of Interaction," in *Distance Education: New Perspectives,* K. Harry, M. John, and D. Keegan (Eds.), Routledge, London, 1993, pp. 12-24.

O'Connor, C., Sceiford, E., Wang, G., and Foucar-Szocki, D. "Departure, Abandonment, and Dropout of E-Learning: Dilemma and Solutions," paper presented at the TechLearn 2003 Conference, Orlando, FL, November 2-5, 2003.

Pask, G. *Conversation, Cognition, and Learning*, Elsevier, New York, 1975.

Phillips, D. C. "The Good, the Bad and the Ugly: The Many Faces of Constructivism," *Educational Researcher* (24:7) 1995, pp. 5-12.

Piaget, J. *Six Psychological Studies*, Vintage, New York, 1967.

Schallert, D. L. "'The Significance of Knowledge': A Synthesis of Research Related to Schema Theory," in *Reading Expository Material,* W. Otto and S. White (Eds.), Academic Press, Inc., New York, 1982, pp. 13-48.

Schank, R. C., and Cleary, C. *Engines for Education*, Lawrence Erlbaum Associates, Mahwah, NJ, 1995.

Voutilainen, A. "Helsinki Taggers and Parsers for English," in *Corpora Calore: Analysis and Techniques in Describing English,* J. M. Kirk (Ed.), Rodopi, Amsterdam, 2000, pp. 319-334..

Zhang, D., and Nunamaker, J. "A Natural Language Approach to Content-Based Video Indexing and Retrieval For Interactive E-Learning," *IEEE Transactions on Multimedia* (6:3), 2004, pp. 450-458.