LEARNING WITH VIRTUAL MENTORS:

HOW TO MAKE E-LEARNING INTERACTIVE AND EFFECTIVE?

by

Jinwei Cao

Copyright © Jinwei Cao 2005

A Dissertation Submitted to the Faculty of the

COMMITTEE ON BUSINESS ADMINISTRATION

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY
WITH A MAJOR IN MANAGEMENT

In the Graduate College

THE UNIVERSITY OF ARIZONA

2005
As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Jinwei Cao entitled Learning With Virtual Mentors: How To Make E-Learning Interactive And Effective and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

__________________________________________ Date:_______________
J. Leon Zhao

Final approval and acceptance of this dissertation is contingent upon the candidate’s submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

__________________________________________ Date:_______________
Dissertation Director: Jay F. Nunamaker, Jr.
STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED: ______JINWEI CAO________________
ACKNOWLEDGEMENTS

Pursuing a Ph.D. degree is the most exciting and memorable experience in my life. It is a tough journey filled with tears and laughter, pain and pleasure, frustration and hope. Without many people’s kind help, I could have never reached the end of this journey. Though it will not be enough to express my gratitude in words to all those people who have helped me, I would like to say:

Thank you to my dissertation advisor, Dr. Jay F. Nunamaker, for patiently supervising me and always guiding me in the right direction. His encouragement, support, and advice have been and will always be a fortune for my academic career.

Thank you to my other dissertation committee members, Dr. Judee K. Burgoon, Dr. Leon Zhao, Dr. Jerzy Rozenblit and Dr. Kevin McNeil, for their valuable suggestions and feedback. Dr. Burgoon has provided very detailed guidance for my dissertation study. What I really learned from her, however, is her attitude to work and life – always aiming for excellence. Dr. Zhao was always available when needed. His encouragement and help made me feel confident to overcome every difficulty I encountered.

Thank you to the research team at the Center for the Management of Information (CMI) for collaboration and support, as well as the Ford Foundation and the U.S. Air force for sponsoring the projects. I have enjoyed every moment I spent with the current and former members of CMI. Special thanks to: Betty Albert, Dr. Dongsong Zhang, Dr. Queen Booker, Xiangyu Zhang, Dr. Janna M. Crews, Ming Lin, Karl Weirs, Chris Steinmeyer, and Amit Deokar.

Thank you to all faculty members at the Department of Management Information Systems for their advice on my PhD study and dissertation research, especially Dr. Olivia Sheng, Dr. Mohan Tanniru, and Dr. Denial Zeng.

Thank you to my fellow PhD students and friends for companying me in this tough journey and sharing my feelings, particularly Fang Chen, Jie Xu, Zan Huang, Yiwen Zhang, Yilu Zhou, Tiantian Qin, and Chris Diller.

Thank you to my dearest husband and soul mate Qianyi Xu, for feeding me with the most delicious food in the world, patting me when I was frustrated, and sharing his intelligence with me all the time. I would never forget that he gave up his well-paid job in China just to be with me together. His love and support is the best gift I have ever had in the past seven years.

Last but not least, I am greatly indebted to my parents, Hongbo Cao and Jie Sun. Even when my father had a heart attack last year, none of them would let me know so as to enable me to concentrate on my research. I owe my every achievement to both of them.
DEDICATION

This dissertation is dedicated to all my family members and friends who have supported me with love and care, with special dedications to:

QIANYI XU

My husband, my lifelong partner, and my love for ever.

I am so happy that we can grow up together.

HONGBO CAO AND JIE SUN

My dad and mom.
You raised me up and taught me how to be a good person.

You are my heroes.
# TABLE OF CONTENTS

LIST OF FIGURES ....................................................................................................... 8  
LIST OF TABLES ......................................................................................................... 9  
ABSTRACT .................................................................................................................. 10  
1. INTRODUCTION ................................................................................................... 11  
  1.1. Background and Motivation .......................................................................... 12  
  1.2. Research Objectives .................................................................................... 14  
  1.3. Overview of Chapters .................................................................................... 16  
2. LITERATURE REVIEW ...................................................................................... 18  
  2.1. Information Technology and Learning ....................................................... 18  
    2.1.1. Terms and Definitions ............................................................................. 18  
    2.1.2. A Framework for Research about Technology Mediated Learning ... 22  
  2.2. Learning Theories and Instructional Strategies ........................................... 25  
    2.2.1. Constructivist Theory ............................................................................. 26  
    2.2.2. Discovery Learning ............................................................................... 27  
    2.2.3. Conversation Theory ............................................................................ 27  
    2.2.4. Interaction in Learning ........................................................................... 28  
  2.3. Technologies for Interactive E-Learning ..................................................... 31  
    2.3.1. Computer Support for Collaborative Learning (CSCL) ....................... 31  
    2.3.2. Hypermedia and Web-based Instruction .............................................. 32  
    2.3.3. Multimedia based Virtual Classroom .................................................. 35  
  2.4. Video-based Question Answering ............................................................... 41  
    2.4.1. Video-based Information Retrieval ....................................................... 41  
    2.4.2. Text-based Question Answering ............................................................. 43  
    2.4.3. Video-based Question Answering .......................................................... 46  
3. RESEARCH PROBLEM: VITALURAL INTERACTION AND  
   LEARNING WITH VIRTUAL MENTORS ...................................................... 49  
  3.1. Adding Interactivity into E-Learning Systems: The LVM Model .............. 50  
  3.2. System Architecture .................................................................................... 57  
    3.2.1. Using PowerPoint Slides to Improve Video-based Question Answering ... 61  
4. A MULTI-METHODOLOGICAL RESEARCH APPROACH .................. 67  
  4.1. System Development .................................................................................... 73  
  4.2. Experimentation ......................................................................................... 75  
  4.3. Observation .................................................................................................. 77
# TABLE OF CONTENTS - CONTINUED

5. EARLY EXPLORATION OF THE LVM MODEL: THE AGENT99 STUDY ........................................................................................................................... 80

5.1. Agent99 Phase I: Build A Virtual Classroom For Deception Detection ................................................................................................................................. 81

5.1.1. System Development .................................................................................................................. 82
5.1.2. Field Experiment for Theory Testing .......................................................................................... 84
5.1.3. Usability Testing ......................................................................................................................... 88

5.2. Agent99 Phase 2: Exploring the Effect of Virtual Interaction ................................................................................................................................. 91

5.2.1. System Development .................................................................................................................. 91
5.2.2. Field Experiment for Theory Testing .......................................................................................... 93
5.2.3. Usability Testing ......................................................................................................................... 97

6. INVESTIGATING THE LVM MODEL: THE LVM STUDY .............................................. 100

6.1. System Development .................................................................................................................. 100
6.2. Experimentation and Observation .............................................................................................. 107

6.2.1. Study Design .............................................................................................................................. 109
6.2.2. Analysis and Results .................................................................................................................. 122

7. CONCLUSION .......................................................................................................................... 173

7.1. Expectations and Outcomes ........................................................................................................ 173
7.2. Implications .................................................................................................................................. 181
7.3. Limitations and Future Research ............................................................................................... 182

APPENDIX A. CODEBOOK FOR OPEN-ENDED RESPONSES ANALYSIS .......................................................................................................................... 185

REFERENCES ........................................................................................................................................ 187
LIST OF FIGURES

Figure 1. Information Technologies and Learning – Key Terms and Their Relationships ..................................................................................................................... 22
Figure 2. A Framework for TML Research ..................................................................................................................... 23
Figure 3. The Interbook Screenshot ..................................................................................................................... 34
Figure 4. IBM Lotus Virtual Classroom ..................................................................................................................... 37
Figure 5. A Real-Time Interactive Virtual Classroom System Developed At the University Of Washington ..................................................................................................................... 37
Figure 6. Stanford Online Courses ..................................................................................................................... 39
Figure 7. E-Classroom in the LBA System (Zhang et al. 2004b) ..................................................................................................................... 40
Figure 8. Question Template from (Zhang et al. 2004a) ..................................................................................................................... 47
Figure 9. The LVM Model ........................................................................................................................................... 52
Figure 10. System Architecture of the LVM Prototype ..................................................................................................................... 58
Figure 11. User Interface of the Prototype System ..................................................................................................................... 60
Figure 12. Three Steps of the QA Process ..................................................................................................................... 62
Figure 13. Multi-methodological Approach to IS research (Nunamaker 1992) ........................................................................................................................................... 69
Figure 14. IS Research Framework (Hevner & March 2003) ..................................................................................................................... 70
Figure 15. A New Vision of the Benefits of Multi-methodological IS Research ........................................................................................................................................... 71
Figure 16. A General Design Research Methodology ..................................................................................................................... 74
Figure 17. Agent99 Trainer I Watch Lecture ..................................................................................................................... 83
Figure 18. Agent99 Trainer I View Example ..................................................................................................................... 84
Figure 19. Agent99 Trainer II Virtual Classroom ..................................................................................................................... 93
Figure 20. Agent99 Trainer II Pop-up Quiz ..................................................................................................................... 93
Figure 21. LVM - Virtual Classroom with Full Interactive Functions ..................................................................................................................... 102
Figure 22. LVM - Answers to Student’s Question ..................................................................................................................... 103
Figure 23. LVM – Pop-up Questions ..................................................................................................................... 104
Figure 24. LVM – Feedback to Pop-up Questions ..................................................................................................................... 104
Figure 25. LVM - Virtual Classroom without Outline and Navigation Buttons ..................................................................................................................... 105
Figure 26. LVM - Virtual Classroom without QA ..................................................................................................................... 106
Figure 27. LVM - Virtual Classroom with Linear Playback ..................................................................................................................... 106
Figure 28. Scree Plot for the MSLQ Items Analysis ..................................................................................................................... 125
Figure 29. Scree Plot for the Post-Experiment Items Analysis (Session 1) ..................................................................................................................... 126
Figure 30. Scree Plot for the Post-Experiment Items Analysis (Session 2) ..................................................................................................................... 126
Figure 31. Means of Interactive Learning Activities in Two Sessions ..................................................................................................................... 128
Figure 32. Means of Test Scores (%) in Both Sessions ..................................................................................................................... 131
Figure 33. Means of Test Scores (%) in Session 2 ..................................................................................................................... 133
Figure 34. Means of Test Scores (%) from the Beginning to the End ..................................................................................................................... 133
Figure 35. Means of Effect in Both Sessions ..................................................................................................................... 136
Figure 36. Means of INTER in Both Sessions ..................................................................................................................... 138
Figure 37. Means of SUS in Both Sessions ..................................................................................................................... 139
Figure 38. Means of LSP in Both Sessions ..................................................................................................................... 140
LIST OF TABLES

Table 1. Evaluation of the QA Module ................................................................. 66
Table 2. Experimental Design of Experiment 1 .................................................. 86
Table 3. Experimental Design for Experiment 2 ............................................... 95
Table 4. Experimental Design ........................................................................... 108
Table 5. Reliabilities for the Constructs in the Post-Experiment Questionnaire .... 123
Table 6. Reliabilities for the Constructs in the Pre-Experiment Survey ............... 123
Table 7. Reliabilities for the Updated Constructs in the Pre-Experiment Survey .... 125
Table 8. Means and Standard Deviations of Interactive Learning Activities .......... 127
Table 9. Means and Standard Deviation of Learning Performance ..................... 130
Table 10. Means and Standard Deviations of Self-Reported Learning Outcomes .... 134
Table 11. Means and Standard Deviation of Learning Motivation/Strategy/Style Measures ................................................................. 141
Table 12. Regression Coefficients of Activity ..................................................... 141
Table 13. Regression Coefficients for Overall Learning Performance Gain .......... 142
Table 14. Regression Coefficient for Learning Performance Gain in Session 2 ...... 142
Table 15. Responses to Questions about System Features ................................. 145
Table 16. Codes Frequencies for Question 1 Responses ..................................... 148
Table 17. Codes Frequencies for Question 2 Responses ..................................... 152
Table 18. Codes Frequencies for Question 3 Responses ..................................... 157
Table 19. Codes Frequencies for Question 4 Responses ..................................... 159
Table 20. Codes Frequencies for Question 5 Responses ..................................... 162
Table 21. Codes Frequencies for Question 6 Responses ..................................... 166
Table 22. Codes Frequencies for Question 7 Responses ..................................... 169
ABSTRACT

This dissertation studies how information technologies, such as automatic question answering (QA), can add interactivity into a multimedia-based e-learning system and turn it into a “virtual mentor” to provide students with interactive, one-on-one instruction. It also explores the key factors of making such a “virtual mentor” as effective as a real mentor.

Based on a review of multiple learning theories and technologies, an exploratory model for studying the effectiveness of interactive e-learning, named “Learning with Virtual Mentors (LVM)”, is proposed and a prototype system is developed to implement the LVM model. A series of studies, including controlled experiments and surveys have been conducted to explore the relationships among the core constructs of the LVM model: learning phases, system interactivity, learner characteristics, learning activity and learning outcomes.

Findings indicate that learning phases and some learner characteristics such as learning style affect students’ behaviors, performance, and perceptions in e-learning partly as we expected. Furthermore, the virtual interaction impacts student behaviors, 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.
1. INTRODUCTION

Show me the army with better trained soldiers and I will show you the victor of the battle. Sun Tsu

Learning is critical for any individual or organization to be successful in the current knowledge-based economy. In recent years, advances of information technology have affected the learning market dramatically. Thousands of online courses, including degree and certificate programs, are now offered by universities and corporations world-wide. For example, in 2001, MIT announced its commitment to make material from virtually all of its courses freely available on the Web for non-commercial use (http://web.mit.edu/newsoffice/nr/2001/ocw.html). The Internet Learning Solutions Group in Cisco Systems Ltd. creates a learning portal containing all of the learning resources that the company had for more than 10,000 systems engineers and account managers in the field, for the employees of the 40,000 channel partners who resell Cisco products, and for hundreds of thousands of end-user customers (Galagan 2001). 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 or 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. It was predicted that U.S. online education student enrollment will top one million by 2005 (Gallagher 2004). According to International Data Corporation (IDC,
http://www.idc.com), the global market for corporate e-learning will be $21 billion in 2008, up from $6.5 billion in 2003 (Hedin et al. 2004).

However, the quick growth of the market demand of e-learning also raises people’s concern of the e-learning quality. Information technology can help people quickly get massive volume of information, but will it guarantee effective learning? Recently, Information Systems (IS) researchers have become more and more interested in studying how to make e-learning effective besides investigating the business benefits of e-learning. This dissertation provides an exploratory view to this “effectiveness” problem by studying the relationships among multimedia, interaction and effectiveness of e-learning, based on a combined understanding of learning theory and information technology.

1.1. Background and Motivation

The effectiveness of e-learning may be improved in many ways. In current e-learning programs, multimedia learning materials such as videotaped lectures and PowerPoint slides are more and more commonly provided as a way to help learners engage in the learning process and learn more effectively. For example, in online courses at Stanford University (http://scpd.stanford.edu/scpd/students/onlineClass.htm), a video of an instructor is provided and synchronized with his/her PowerPoint slides. Multimedia lectures are considered to be able to give students a perception of listening to a lecture in real-time, make them pay more attention to the learning task, and help them retain more information through vivid and rich presentations (Agius & Angelides 1999). It seems that online multimedia lectures can give students a perception of learning with a private teacher (a.k.a. mentor) because the students can watch the lecture alone and feel that the
instructor is talking to just him or herself. However, simply watching an instructor talking in a lecture video is still quite different from learning with a real mentor. An important factor of learning, “interaction,” is usually missing in multimedia online lectures (O’Connor et al. 2003). For example, when a student learns from a real mentor, the student can ask questions at any time, and the mentor can ask the student questions and assess the student’s knowledge based on his or her answers and thus provide customized instructions, which may trigger more questions from the student. Many modern learning theories indicate that effective learning requires such an iterative interaction process between the learner and the knowledge providers (Bruner 1960; Bruner 1966; Pask 1975); however, it cannot be realized by just a linear playback of the video of instructions.

Although collaborative learning technologies such as chat rooms or discussion forums can be added 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 need an answer right away (e.g., for an instant task required on the job), the mentor recorded in the videos may not be accessible either face-to-face or online at that certain time. In this situation, students have to find the answer by themselves, either by sequentially looking through the lecture videos or searching on the Web. However, nobody is patient enough to look through a 60 minutes long, linear and unstructured video; while Web searching may not get authenticated answers as those directly from the mentor.

So how can we integrate interaction into multimedia e-learning with minimal reliance on human mentors? Can we use advanced information technologies to turn a “dumb”
repository of unstructured video instructions into a “live” virtual mentor that learners can directly interact with at any time? What activities are associated with the process of learning with such a virtual mentor? Do learners behave differently in learning with virtual mentors as compared to learning with real mentors? Will this type of “virtual” interaction make e-learning more effective? What factors influence the effectiveness of learning with virtual mentors? How should such a system be deployed to let students achieve better learning performance?

1.2. Research Objectives

Inspired by the above questions, this dissertation studies whether information technologies, such as automatic question answering (QA), can add interactivity into a multimedia-based e-learning system and turn it into a “virtual mentor” to provide students with interactive, one-on-one instruction. Here a virtual mentor (VM) is defined as “a virtual, multimedia representation of a knowledgeable mentor that a learner can interact with and learn from, through activities simulating the normal interactive activities between the learner and a human mentor, such as personalized instruction, questioning and answering, as well as assessment and feedback.” Such simulation of interactive activities is referred to as “virtual interaction” in this dissertation.

In addition and more importantly, this dissertation studies how this type of “virtual interaction” affects the effectiveness of e-learning. Following a framework for technology mediated learning research (Alavi & Leidner 2001), this dissertation explores the key factors of making such a virtual mentor effective by looking at both information technologies and learning theories, as well as studying students’ psychological learning
processes. An exploratory model of these factors for the effectiveness of interactive e-learning, named “Learning with Virtual Mentors (LVM),” is proposed based on multiple learning theories. The primary goal of this dissertation is therefore to validate and modify the LVM model with a series of research studies.

An important feature of this dissertation study is the triangulation of different research approaches such as system development, experimentation, survey and interview, as well as different research disciplines such as information technology, instructional strategy, and psychology. A prototype system has been developed based on the proposed LVM model. Besides providing students multimedia instructions in a structured way and giving them assessment and feedback, the prototype system offers a unique Question Answering (QA) module which allows students asking a natural language question and receiving answers from archived multimedia instructions. A series of studies, including controlled experiments, closed-ended and open-ended surveys, and semi-structured interviews have been conducted to explore the relationship among the learning technologies, instructional strategies, students’ mental models, their learning behaviors, and their learning effectiveness. The integration of so many different perspectives provides us a comprehensive understanding of the problem. We expect that the findings of this dissertation study will provide theoretical, strategic, and practical guidance to e-learning researchers and developers, in terms of how to improve interactivity and effectiveness of e-learning.
1.3. Overview of Chapters

This section provides an overview of each of the remaining chapters in this dissertation. The research conducted is reported herein, as follows:

- **Chapter 2. LITERATURE REVIEW** reviews literature on topics related to 1) *Information Technology and Learning*, such as Technology Mediated Learning, E-Learning, and Virtual Learning Systems; 2) *Learning Theories and Instructional Strategies*, including Constructivist Theory, Discovery Learning, and Conversation Theory, and Interaction in Learning; 3) *Technologies for Interactive E-Learning*, such as Computer Support for Collaborative Learning (CSCL), Hypermedia and Web-based Learning Systems, and Multimedia-based Virtual Classroom; 4) *Video-based Question Answering*, including Video-based Information Retrieval, Text-based Question Answering, and Video-based Question Answering.

- **Chapter 3. RESEARCH PROBLEM: VIRTUAL INTERACTION AND LEARNING WITH VIRTUAL MENTORS** describes the LVM model and the resulted prototype system. It focuses on describing the research problem, research questions, system architecture, and the key technology innovations in the prototype system.

- **Chapter 4. A MULTI-METHODOLOGICAL RESEARCH STUDY** discusses the rationale and benefits of using a multi-methodological approach in this dissertation study. It also describes the research methods selected, including system development, experimentation, and observation.
• Chapter 5. EARLY EXPLORATION OF THE LVM MODEL: THE AGENT99 STUDY describes an empirical study named the Agent99 Trainer (A99) study that helped us form the LVM model.

• Chapter 6. INVESTIGATING THE LVM MODEL: THE LVM STUDY presents the specific study designs for the LVM study, including measurements and instruments. It also describes the analysis of the data collected, and presents both quantitative results and qualitative results from the different research methods, as well as their triangulation.

• Chapter 7. CONCLUSION summarizes the research expectations and outcomes, contributions, limitations, and the opportunities for future research.
2. LITERATURE REVIEW

2.1. Information Technology and Learning

It is not a new idea at all that information technologies (IT), including technologies such as computing, communication, data management, and so on, can be used to support learning. Ever since the computer was invented, people had tried many ways to use it to either enhance classroom teaching and learning or directly deliver instructions. The dramatic increase in the practice and research of IT in teaching and learning in the past decades has brought in many new concepts. Several terms have been attached to the innovation and creation that has been occurring in this evolution era. However, many terms do not have solidified definitions and have been used disorderly. Therefore, in this section, we first review and summarize some key terms related to the application of IT in learning, so that we can term and describe our research without misunderstanding. We then review a framework on how to study the impact of IT in learning, upon which this dissertation research is developed.

2.1.1. Terms and Definitions

TECHNOLOGY MEDIATED LEARNING (TML) is defined as “an environment in which the learner’s interactions with learning materials (readings, assignments, exercises, etc.), peers, and/or instructors are mediated through advanced information technologies.” (Alavi & Leidner 2001) This term summarizes all types of application of IT in learning, including both using IT in traditional classrooms (e.g., instructor console or stand-alone
student computers in a tech-classroom (Leidner & Jarvenpaa 1995)) and using IT as a
directly learning delivery tool.

DISTANCE LEARNING is defined by the United States Distance Learning
Association as "the acquisition of knowledge and skills through mediated information
and instruction, encompassing all technologies and other forms of learning at a distance"
(http://www.usdla.org/html/resources/dictionary.htm#d). This definition stresses the role
of distance, and any technology including more traditional delivery methods, such as the
postal service can be used to breech the distance. Therefore, TML may be either distance
or local, while distance learning may be mediated by information technology or not.

ELECTRONIC LEARNING (E-LEARNING) is a de facto term that has been used
frequently in both practice and research. There are many different definitions of e-
learning and no solidified definition so far. The most frequently cited definition of e-
learning, however, is a “Term covering a wide set of applications and processes, such as
Web-based learning, computer-based learning, virtual classrooms, and digital
collaboration. It includes the delivery of content via Internet, intranet/extranet
(LAN/WAN), audio- and videotape, satellite broadcast, interactive TV, CD-ROM, and
more.” (ANTA 2003); Learning Circuits http://www.learningcircuits.org/glossary). This
well accepted definition and its variations emphasize e-learning as the electronic delivery
methods of teaching and learning. Therefore in this dissertation we view e-learning as a
type of TML, but not necessarily distance learning. The instructor and students can be at
the same place while all instructions are delivered by an e-learning system. In this case
the instructor’s role is simply changed to a facilitator. In this dissertation, all TML
applications that are not direct support for traditional, face-to-face, instructor-led classroom learning (e.g. tech-classroom) are classified under the umbrella of e-learning.

COMPUTER-ASSISTED INSTRUCTION (CAI)/COMPUTER-BASED TRAINING (CBT) are also two commonly used terms in describing the applications of IT in learning. Development and study of CAI started as early as in the 1960s at Stanford University (Suppes & Morningstar 1969). Early attempts of CAI comprised interactive software programs providing learner feedback, lesson branching, and student record keeping. Although successful for some types of learning, most CAI programs’ ability to provide adaptive instruction was limited to branching between static screens. While the term CAI has been used in educational settings, the term CBT has been commonly used to refer to the similar systems used in organizational training. (Leidner & Jarvenpaa 1995)

Therefore, the terms CBT and CAI can be used interchangeably. In this dissertation, we view CAI/CBT as a type of e-learning in which the electronic media only refers to computers. In addition it implies a more planned approach for learning, i.e., instruction.

WEB-BASED INSTRUCTION (WBI), also called WEB-BASED TRAINING (WBT) is a more advanced type of CBT. It is defined as a hypermedia-based instructional program which utilizes the attributes and resources of the World Wide Web to create a meaningful learning environment where learning is fostered and supported (Kahn 1997). In WBT, individualized instruction is delivered over public or private computer networks and displayed by a Web browser. Therefore WBT can be updated very rapidly, and access to training can be controlled by the training provider.
While CAI/CBT focus more on instruction, a more planned approach for learning, COMPUTER SUPPORTED COLLABORATIVE LEARNING (CSCL) focuses on “the use of information and communications technology (ICT) as a mediational tool within collaborative methods (e.g. peer learning and tutoring, reciprocal teaching, project- or problem-based learning, simulations, games) of learning” (Wasson 1998). CSCL research is grounded on very different concepts of learning, pedagogy, research methodology, and research questions than CAI/CBT. It is a type of e-learning that focuses on collaborative learning.

Similarly, an ASYNCHRONOUS LEARNING NETWORK (ALN) is “a teaching and learning environment located within a CMC system designed for anytime/anyplace use through computer networks. ALNs consist of a set of group communication and work "spaces" and facilities constructed in software. They are virtual facilities for interaction among the members of a class, rather than physical spaces” (Hiltz & Wellman 1997). ALN also focuses on collaborative learning but it is only one type of CSCL, because it only emphasizes on asynchronous collaboration.

These key terms and their relationships are illustrated in Figure 1. From the definitions discussed above, we can see that students using CAI/CBT systems directly interact with computers while students using CSCL/ALN systems interact with live people via electronic media. One of the goals of this dissertation, however, is to build a special type of interactive e-learning system in which students still interact with a computer but have the feeling of interacting with live people. In other words, we want to
provide students a “virtual mentor” – a virtual interactive e-learning system (illustrated by the dotted circle in Figure 1).

2.1.2. A Framework for Research about Technology Mediated Learning

As stated earlier, research about Technology Mediated Learning (TML) started as early as in the 1960’s. Researchers have studied many different problems in this area, from developing novel systems to utilizing learning theories. However, most of the existing studies have focused on the direct impact of specific technology features on learning outcomes. A guideline for theoretically grounded and rigorous research of TML is missing. Until 2001, Alavi and Leidner proposed a framework for TML research (see Figure 2), and recommended attention to a greater depth of research questions in the area of TML, especially, “How can technology enhance learning?” (Alavi & Leidner 2001).
Depicted in Figure 2, this framework suggests TML researchers to pay attention to relationships among technology features and the relevant instructional, psychological, and environmental factors that will enhance learning outcomes.

There are different types of **learning outcomes** that may be of interest to TML researchers. In (Alavi & Leidner 2001), learners’ intellectual skills, motor skills, verbal information, cognitive strategy, attitude, affective reactions to TML (e.g., satisfaction with the TML experience) and efficiency (cost and learning time) of TML environments have been given as examples of learning outcomes. However, many of these learning outcomes about knowledge and capability are not directly observable. Alavi and Leidner thus suggest observing and measuring the action and performance that result from learning.

In a given learning context, **psychological learning processes** refer to states within the learner that are involved in learning, including the learner’s cognitive and information processing activities, motivation, interest, and cognitive structures (memory) (Alavi &
Leidner 2001). Without affecting these processes first, technology features cannot directly affect learning outcomes.

As stated earlier, most existing TML research focuses on studying the information technology features, such as collaborative technologies, media synchronization, and so on. Technology features can influence the learning outcomes through the direct support of the underlying psychological processes (e.g., facilitating cognitive information processing activities such as search, scanning, transformation, or comparison of information) (Alavi & Leidner 2001).

**Instructional strategy**, on the other hand, refers to methods and models for presenting (selection and display mode and format), sequencing (order of “topics” to be presented), and synthesizing (establishing relationships among topics) subject-matter content. Instructional strategies aim to facilitate learning by activating the psychological processes required by learners for achieving the desired learning outcomes (Alavi & Leidner 2001). Information technology can also enhance the execution of instructional strategies (e.g., providing new forms of content presentations) and then facilitate the activation of learners’ psychological processes. This finally leads to more effective or efficient learning outcomes.

To summarize, this framework suggest that information technology features can enhance learning outcomes by facilitating a rich instructional strategy and by directly eliciting certain cognitive information processing activities of the learner (Alavi & Leidner 2001). Therefore, TML research studies should take into account all aspects of
2.2. Learning Theories and Instructional Strategies

As the framework of TML research indicates, instructional strategies can significantly affect the effectiveness of information technology supported learning. The instructional strategies, on the other hand, are mostly drawn from learning theories.

Modern learning theories can be classified into two major divisions. The first division can be represented by theories such as the Behaviorism (Skinner 1938; Skinner 1971), the Contiguity Model (Guthrie 1935), and the Stochastic Model (Estes 1950). The center point of these theories is a stimulus-response model of learning. They view learners as passive recipients of information sent by the expert, and that learning is a consequence of association between particular stimuli (e.g. instructor-controlled steps and the positive and negative feedback given to the learner) and responses. This division of learning theories fit closely with traditional lecture-based learning, where instructor-controlled teaching is followed by tests and quizzes designed to evaluate the correct conceptualization and retention of course material (Skinner 1971).

The behaviorism learning theories were often criticized for ignoring the humanity and the individuality of learners. To address these limitations, another division of learning theories was developed since the 1960’s. Represented by the constructivist theory (Bruner 1966), these theories view learners as active explorers at the center of the learning process. Particularly, 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). These theories form the theoretical foundation for this dissertation study. Therefore they are described in detail in the following sections.

2.2.1. Constructivist Theory

The first of these three theories, the constructivist theory, is a learning theory that is broadly applied in both traditional type and computer-supported learning. Constructivist theory claims that learning is an active process in which learners construct new ideas or concepts based upon their current/past knowledge (Bruner 1966). Constructivist theory emphasizes that learning should be “learner-centered” (Phillips 1995) instead of “teacher-centered” in traditional lecture-based learning. Students are active agents engaging in their own knowledge construction based on personal experience and interpretation (Merrill 1991). As far as instruction is concerned, the instructor should try and encourage students to discover principles by themselves. Guidance and concrete teaching are only provided whenever necessary. The instructor and student should engage in an active dialog (i.e., Socratic learning). The task of the instructor is to translate information to be learned into a format appropriate to the learner's current state of understanding. Curriculum should be organized in a spiral manner so that the student continually builds upon what they have already learned. Such a student-centered guided learning environment is considered more appropriate for ill-structured domains or higher-level learning ((CTGV) 1991), which are also usually features of adult learning or on-the-job learning.
2.2.2. Discovery Learning

The constructivist theory was actually developed from an earlier theory called discovery learning. Discovery learning is "an approach to instruction through which students interact with their environment--by exploring and manipulating objects, wrestling with questions and controversies, or performing experiments" (Ormrod 1995). In discovery learning, students are provided with data and are called upon to question, explore, or experiment. By questioning the teacher, they are expected to ascertain the particular principle hidden in the lesson objective. It is believed that they will be better able to remember and apply what they have learned in this way (Bruner 1960).

The key propositions of the discovery learning theory and constructivist theory are similar and overlapped to some extent, because they both emphasize that to learn something better, students need to discover principles by themselves through questioning the teacher and engage in active dialogs. However, constructivist theory also emphasizes the importance of past knowledge and feedback.

2.2.3. Conversation Theory

The conversation theory, on the other hand, originates from a cybernetics framework and attempts to explain learning in both living organisms and machines (Pask 1975). The fundamental idea of this theory is that learning occurs through conversations about a subject matter, and the conversations serve to make knowledge explicit. Conversations can be conducted at a number of different levels: natural language (general discussion), object languages (for discussing the subject matter), and metalanguages (for talking about learning/language). In human’s learning process, the natural language conversations such
as questioning and answering are the major level of conversation, and they enable learning by exploration and reflection (Schank & Cleary. 1995). Particularly, once a question has been generated by memory, memory is set to learn since it knows where to place any answer it finds. Also asking questions leads learners to think more deeply about their experiences and build more detailed indices between their experiences and the explicit explanation they received. Therefore, asking questions, receiving explanations, and the generation of self-explanations are critical if students are to learn anything more than isolated facts.

2.2.4. Interaction in Learning

The three modern learning theories described above all indicate that high quality learning must incorporate active interactions initiated by students, especially a questioning and answering process. Therefore, the instructional strategies for an effective e-learning system must include strategies for facilitating active interactions in learning.

There is a growing body of research investigates the strategies and effects of interactions in learning, especially interactions in the TML environment.

Basically, interactions in learning can be classified as the following three types (Moore 1993):

1) LEARNER – CONTENT INTERACTION: it is usually a static interaction between the student and static learning contents. “Traditional” activities of this type of interaction include questioning the training materials. This type of interaction was least studied because people usually assume that one can only interact with a human. However, it is said that interaction occurs not only with the originator of knowledge but also with
knowledge itself, represented in media (tutors, peers, technologies, real or virtual objects and entities) as information (Bates 1995). Therefore, media of knowledge should not be an obstacle for successful interaction. In addition, many research studies show that thoughts can be triggered, meaning can be made and schemata can be created and recreated through interacting with certain piece/s of information received and revisited (Anderson & Pearson 1984; Piaget 1967; Schallert 1982).

2) LEARNER(S) - TUTOR(S) INTERACTION, and 3) LEARNER(S) - LEARNER(S) INTERACTION: these two types of interaction are among humans, and they are the interaction forms that people are most familiar with. Therefore, most research studies are focusing on these two types of interaction, especially in the research of Computer Supported Collaborative Learning (CSCL). According to (Hiltz & Wellman 1997), if collaboration rather than individual learning designs were used in an online class, students should be more motivated to actively participate and should perceive the medium as relatively friendly and personal as a result of the online social interactions. This increased active group interaction and participation in the online course, hence, resulted in higher perceptions of self-reported learning. Whereas individuals working alone online tended to be less motivated, perceive lower levels of learning, and score lower on the test of mastery.

In CSCL, researchers usually distinguish two types of interactions between learner-tutor or learner-learner. The first one, **synchronous** interaction, requires that all participants of interaction are online at the same time. Examples include Internet voice telephone, video teleconferencing, text-based chat systems, instant messaging systems,
text-based virtual learning environments, graphical virtual reality environments, and net-based virtual auditorium or lecture room systems. Synchronous interaction promotes faster problem solving, scheduling and decision making, and provides increased opportunities for developing affect (Moore 1993). In 2000, Hron et al. studied the interaction in virtual learning groups supported by synchronous communication. They found that learning in virtual environments can be greatly enhanced by content-related dialogues with minor off-task talk, coherent subject matter discussion with explanation, and equal participation of students supported by synchronous interaction (Hron et al. 2000). However, the cost of synchronous interaction is usually very high, and synchronous interaction is more constricted due to time differences.

The second one is asynchronous interaction, in which learners or tutors have freedom of time and location to participate in the interaction, examples including interaction using E-Mail, ListServ type mail systems, discussion forums, and bulletin board systems. It has been reported that by extending interactions to times outside of classes, more persistent interaction and closer interpersonal bonds among students can occur (Haythornthwaite 1999). Thus, while one cannot totally simulate a real classroom with synchronous interaction, one can offer asynchronous interaction that provides time for better reflection, and allows global communication un-bounded by time zone constraints. Asynchronous interaction thus is more commonly provided in CSCL systems than the costly synchronous interaction (Moore 1993).

Although the last two types of interactions have been demonstrated as good strategies for supporting effective learning, they all have to rely on the availability of live people.
Particularly for the interaction between learners and tutors, even though asynchronous interaction can temporarily relieve the problem when the instructor is not available, it does prevent students from receiving the feedback in time, which is undesirable in many situations. Therefore, in this dissertation, we focus on finding technologies to support a special type of Learner – Content interaction (virtual interaction), which does not rely on the availability of live instructors but has the advantages of both synchronous and asynchronous Learner-Tutor interactions. However, research is lacking on this mode of interaction in learning. We thus review the existing technologies for interactive e-learning, with the intention of finding the most appropriate technology for implementing such virtual interaction.

2.3. Technologies for Interactive E-Learning

In this section, technologies that have been used to support different types of interactions in e-learning systems are reviewed and summarized.

2.3.1. Computer Support for Collaborative Learning (CSCL)

The most studied interactive e-learning technology is often referred to as Computer Support for Collaborative Learning (CSCL). As described earlier, CSCL refers to the use of computers to support interaction among peers or with instructors, and to promote shared experience and/or mental models. In CSCL systems, learning occurs from group members’ exposure to each other’s thinking, opinions, and beliefs, as well as obtaining and providing feedback for clarification and comprehension. Therefore the key for effective learning in CSCL is the active participation and interaction of learners.
Many CSCL systems support the interaction through the use of communication technologies such as e-mail and on-line chat facilities. In addition, more and more CSCL systems begin to offer Web-based software that simulates many collaboration activities that take place in classroom. Two types of software that are believed to be the key for creating active participation and interaction are: 1) quiz routines designed for self-testing and providing feedback, and 2) computer conferencing systems that support synchronous or asynchronous on-line discussions for exchanging ideas and information (Hiltz & Turoff 2002). Examples of CSCL systems include: the early trial of using Group Support Systems (GSS) in classroom settings (Walsh et al. 1996), the Knowledge Forum system developed by the Learning in Motion Ltd. (http://www.learn.motion.com/) that allows users to create a knowledge-building community in which they can share notes and connect ideas, as well as the Asynchronous Learning Networks (ALN) that connect the learners and the instructors via the Internet (Hiltz & Wellman 1997).

However, as discussed earlier, the communication and conferencing technologies used in CSCL can only support interactions among live people. Therefore these technologies are not germane to this dissertation.

2.3.2. Hypermedia and Web-based Instruction

The most common technology that has been used to implement the Learner-Content type of interaction is hypermedia. Hypermedia, which is a combination of multimedia and hypertext, has a node-and-link structure inherent to the organization of information and is based on integrated media (Heller 1992). Students interact with hypermedia systems by exploring the links to multimedia learning materials. Hypermedia can be
easily implemented in a Web-based Instruction (WBI) system and be delivered to
students through the Internet. In such WBI systems, students have great flexibility in
choosing the time, pace, frequency and form of learning activities.

Hypermedia technology allows learners to explore a topic in multiple ways through
different node and link structures. Such exploratory interaction between the learner and
the learning content can help the learner create associational links within and across text,
images, video, and other media, and finally build an integrated conceptual model (Kozma
1994). Students are not passive but proactive in interpreting and constructing new
knowledge from the hypermedia information by processing and filtering it through their
existing cognitive structures. In addition, research has shown that multimedia can both
improve students’ higher-level of comprehension of knowledge (Large et al. 1994),
capture students’ interest, and generate subjective feelings of better learning (Demetriadis
et al. 2003). Particularly as reported in (Demetriadis et al. 2003), highly animated media
such as video or audio, can make students feel as if they were in the classroom listening
to the teacher, which creates a more vivid social context of learning and interaction. In
this sense, hypermedia systems can make students’ interaction with contents more vivid
and thus can make students learn more efficiently.

A more advanced type of hypermedia systems is Adaptive Hypermedia Systems
(AHS). As defined by Brusilovsky, "adaptive hypermedia systems are all hypertext and
hypermedia systems that reflect some features of the user in the user model and that
apply this model to adapt various visible aspects of the system to the user" (Brusilovsky
1996). AHS use a domain model to contain a set of domain concepts. The concepts in the
domain model are related to each other, thus forming a kind of semantic network. For each concept in the domain model, an individual user model called an overlay model stores some value as an estimation of the user knowledge level on this concept. Based on this user model, AHS can provide adaptive presentation and navigation to individual students (Brusilovsky 1996). For example, in an AHS application named InterBook (Brusilovsky et al. 1998), links are marked with different colors based on the individual student’s learning model (Figure 3). A red dot means the student is missing some required foreknowledge for understanding the knowledge that is linked. Green dots indicate recommended pages. A white dot indicates that the user already knows the concept(s) explained on this page. In the example, the user is ready to jump to section 1.5 but is still lacking some prerequisite knowledge for subsections of 1.1.

![Interbook Screenshot](image)

**Figure 3. The Interbook Screenshot**

AHS applications such as Inter-Book allow students to be “understood” by the computer system and thus can provide more efficient navigation control (Brusilovsky et al. 1998). However, these systems do not have the function to let users input their questions and only allow them to go along the existing links. Therefore, the interaction between users and system is still limited to some extent.
2.3.3. Multimedia-based Virtual Classroom

Recently, more and more online course providers are claiming that they provide “Virtual Classrooms” for more effective learning. But what is a “Virtual Classroom”? There has not been a consensus on this issue yet.

The Virtual Classroom™ has been registered as NJIT's trademarked name for versions of its Electronic Information Exchange System (EIES2) with special software structures designed to support collaborative learning, including structures to order the transcript of discussions, to force active participation, to allocate unique assignment topics, and an electronic gradebook (Hiltz & Wellman 1997). This system’s major feature, however, is still the asynchronous collaboration. Students are considered learning in a Virtual Classroom when they can participate in class discussions with peers and/or instructors asynchronously or “anytime/anywhere”. The system stores the discussion entries in a permanent, ordered transcript which keeps the equivalent of "bookmarks" to separate anything that is "new" for each individual from items that have already been seen (Hiltz & Wellman 1997). In this sense, the NJIT’s Virtual Classroom™ is more like a computer-mediated communication channel for learning. The lecture type of material still has to be delivered either in face-to-face meetings, or using print media, or via audio/audio-graphic media or CD ROM or other special PC based software.

The most recent use of the name “Virtual Classroom,” however, is from a different perspective. As stated in (Deshpande & Hwang 2001), a virtual classroom environment aims at simulating a real classroom for remote participants, who can receive a live class
feed and are also able to interact and participate in the class by asking questions. This concept is fashioned after a lecture-style classroom environment. Unlike the NJIT Virtual Classroom™, students not only have access to people (peers or instructors) for discussion, but also have access to lecture information that simulates what is provided in a real classroom. The emphasis of this type of Virtual Classroom system, in our understanding, is the multimedia simulation of lecture presentations. In this dissertation, we interpret the Virtual Classroom concept from this perspective.

Considering Virtual Classroom (VC) as the multimedia simulation of lecture presentations, there are many examples available from both commercial applications and academic experimentations, such as the IBM Lotus® Virtual Classroom (Figure 4) and the real-time interactive virtual classroom multimedia distance learning system developed at the University of Washington (Figure 5) (Deshpande & Hwang 2001).

The key technology feature of this type of VC is the integration of multiple media channels for simulating a real classroom presentation. Currently, the most applied method is to use streaming video and audio to capture the class or lecture talk, and use images and texts to represent the presentation materials including electronic slides, written text on a white board, or lecture notes. Some VC applications also have the function for synchronizing these media channels.
Figure 4. IBM Lotus Virtual Classroom

Figure 5. A Real-Time Interactive Virtual Classroom System Developed At the University Of Washington
The VC applications illustrated in Figure 4 and Figure 5 use synchronous communication technologies, particularly videoconferencing technologies to enable real-time interactive two-way or multi-way communications among students and the instructor. Other synchronous communication technologies such as telephone or on-line chat facilities and key response pads have also been used in similar VC environment. As stated earlier, such synchronous interaction provides the instructor with useful feedback to gauge students’ comprehension, and thus allows the instructor to adjust the presentation of material accordingly in real time. However, learning in synchronous VC is seriously constricted to time differences.

Another type of VC applications, which in most cases was not originally named as Virtual Classroom, also simulates real classroom presentations but in an asynchronous way. In these systems, classes or lectures are pre-recorded in a real classroom or in a studio, and then the video/audio of class talk as well as the images/texts of the presentation materials is presented in a similar integrated user interface. The multimedia recordings of the lectures are usually processed so that hyperlinks can be added to help students navigate through different segments of the lecture. This can be viewed as an extension of the hypermedia system. However, there is no guarantee that the instructor will be available for questions when students retrieving the lecture on their convenience. Therefore, the interaction has to be provided using asynchronous communication technologies such as e-mail or online discussion forum. The use of asynchronous communication allows instructor to answer questions beyond the scheduled class period.
One example for this type of VC applications is the system for the online courses provided at Stanford University (Figure 6).

A similar system named as E-Classroom has been provided as part of a Web-based learning system Learning By Asking (LBA; the ancestor of this dissertation study; see Figure 7).
In the E-Classroom, the lecture is segmented into small pieces according to topic transition, and hyperlinks are added to help students navigate through different segments. Therefore, students can interact with the learning content as interacting in a hypermedia system. However, different from the other commonly available VC systems, the LBA system attempts to add in a new type of interaction by allowing students to type in a natural language question and receiving answers extracted from the recorded lectures (Zhang 2004). Although the impact of such interaction on the effectiveness of the learning system was not studied in the LBA study, we believe this new type of interaction is a good bridge between the Learner-Content type of interaction and the Learner-Instructor type of interaction, because we believe the questioning and answering process with the help of video and audio of the instructor can well simulate the live communication between people. Therefore in this dissertation, we select the video-based
question answering technology proposed in (Zhang 2004) to implement the virtual interaction and study its impact on the effectiveness of e-learning.

2.4. Video-based Question Answering

The questioning and answering process we described above can also be viewed as a process of finding the specific video segment that is most relevant to the student’s question. The technology for finding specific segments from a video actually stems from video-based information retrieval (IR). However, to understand the student’s question, natural language understanding technology is also required. The combination of IR and natural language processing is referred to as Question Answering (QA) technology. In this section we first review the actively studied text-based QA technology and then describe in detail the adaptation of it into the video application.

2.4.1. Video-based Information Retrieval

Defined in (Baeza-Yates & Ribeiro-Neto 1999), Information Retrieval (IR) deals with the representation, storage, organization of, and access to the information items in which the user is interested. The general retrieval process can be summarized as the following sub-processes. First, a text database is set up to store the original documents (e.g. Web pages). An index of the text is then built to allow fast searching over large volumes of data. To initiate the retrieval, the user first specifies his or her need which is then parsed and transformed to a query. The query is then processed to obtain the retrieved documents. Finally, the retrieved documents are ranked according to the likelihood of relevance.
The traditional keyword-based information retrieval is used in most current search engines, where users are required to enter keywords to query the server database. The server then searches the index and retrieves the documents that match the keywords entered by the user (Baeza-Yates & Ribeiro-Neto 1999). Keyword search can be easily developed or implemented by using commercial search engine APIs such as Google™ (http://www.google.com/apis/). It has also been used in e-learning systems, for example, the Electronic Campus (SREB 2004). However, this approach is usually inefficient because it relies on the assumption that users can express their needs in terms of keywords accurately. If the query keywords are poorly structured (e.g., problems with synonymy), the search may not work as the user expected or may even return nothing (Quah et al. 2002).

Although traditional information retrieval is based on pure text, multimedia information retrieval, such as retrieving images or video is becoming more and more important because of the quick development of multimedia technologies and growing volumes of multimedia documents. Basically, multimedia information retrieval can be classified based on two different approaches (Brown et al. 2001). Expression-based techniques rely on an example or a physical description of the information that is sought; while semantic approaches rely on the actual content of the media. There are also different types of queries in multimedia information retrieval. Users can either query by example, such as a similar image, or formulate a query from the data representation, e.g., the physical description of the required information such as the textual query to search an
audio or video database. The ultimate goal in multimedia retrieval is to achieve semantic retrieval using fully automated techniques.

Video retrieval is one area of multimedia information retrieval. Similarly, it can also be done by two different approaches. The first one relies on image-retrieval techniques to retrieve video based on key frames extracted from the video. It has not yet become popular, because in most practical situations the user does not have such an image available to formulate the query, and the state of the art in content-based image retrieval has not yet reached the semantic level desired by most users. The second approach is actually more commonly applied currently. This approach searches the audio transcript of the video using the familiar metaphor of free text search. Usually speech recognition is applied to the audio track, and a time-aligned transcript is generated, then the indexed transcript provides direct access to the semantic information in the video. However, speech recognition accuracy can greatly affect the retrieval results, making video retrieval a challenging task. Phonetic retrieval (“sounds-like” retrieval) is a possible complementary method to solve this problem (Brown et al. 2001).

One drawback of the current multimedia information retrieval technology is that most of the systems do not have natural language processing functionalities and therefore can only allow keyword-based searches.

2.4.2. Text-based Question Answering

A combination of IR and natural language processing, automated Question Answering (QA) is the technology that locates, extracts, and represents a specific answer to a user question expressed in natural language. Literally dozens of systems have been
developed for the Q&A track at the Text Retrieval Conference (TREC) since 1999 (Voorhees 1999; Voorhees 2000; Voorhees 2001), but all of them only work with text documents. A text-based QA system would take as input a question such as “What is mad cow disease?” and, instead of extracting a list of relevant documents from the document archive as a keyword-based search engine would do, it produces the answer in specific sentences, e.g. “Mad cow disease is a fatal disease of cattle that affects the central nervous system. It causes staggering and agitation.”

Modern Question Answering (QA) technologies rely on many components, including document retrieval, semantic analysis, syntactic parsing and explanation generation. QA promises an important new way of information access for many types of users (including novices in technology) beyond the keyword query and document retrieval characteristic of today’s information quests, such as those provided by current web search engines. Evaluating TREC question answering (Voorhees 1999) is the motivating force behind a recent surge in question answering research. Systems participating in TREC have to identify exact answers to “factual questions” (who, when, where, what, etc.). Most of TREC QA systems are designed based on techniques from natural language processing, information retrieval and computational linguistics. For example, Falcon (Harabagiu et al. 2000), one of the most successful QA systems, is based on a pre-built hierarchy of dozens of semantic types of expected answers, complete syntactic parsing of all potential answer sources, and automated theorem proving to identify the answers. Such a natural language based QA approach usually follows the following three steps (Voorhees 1999):
1) QUESTION UNDERSTANDING: Derive the answer type from the question, e.g. “Who is” wants a person, and turn the question to a query.

2) DOCUMENT FILTERING: Use keywords in the question as query and use some query expansion (e.g. morphological or lexical, using WordNet, etc.) to do Boolean IR. Only the matched documents will be processed to extract answers.

3) ANSWER EXTRACTION: Perform a shallow parse of the returned documents to detect entities of the same type as the answer. Calculate the relevancy between the processed query and sentences in the documents extracted in step 2. Extract a list of answers ranked by their relevancy to the query.

However, different natural language parsers and relevancy metrics used in the different systems cause performances to vary.

In addition to “deep” linguistic approaches, QA researchers have successfully explored more “shallow” approaches grounded on pattern matching technologies used earlier for information extraction (Soubbotin & Soubbotin 2002). Pattern matching systems performed well in recent TREC QA competitions: the system from InsightSoft (Soubbotin & Soubbotin 2002) won 1st place in 2002 and 2nd place in 2001. Roussinov Robles (Roussinov & Robles 2004) combined the redundancy based approach suggested in (Dumais et al. 2002) with the automatically learned patterns suggested in (Ravichandran & Hovy 2002) into a single probabilistic trainable approach. Pattern-matching QA has advantages over other techniques because it does not require any manually developed rules or substantial linguistic resources. However, pattern-matching QA systems perform well only when there is a large set of documents with redundancy
(such as very large digital libraries, large intranets or the entire WWW). It has fewer advantages when the domain is smaller (Cao et al. 2004c).

2.4.3. Video-based Question Answering

As mentioned earlier, in the LBA study (the ancestor of this dissertation study), Zhang and his colleagues proposed a method that applies the text-based QA approaches to video applications (Zhang & Nunamaker 2004). To extract answers from video lectures according to a student’s question, a natural language based QA approach is applied to a collection of transcribed videos.

In that approach, the speech in video lectures is first transcribed into text manually or automatically using speech recognition software. Each long video of a lecture is also manually segmented into small pieces. The transcripts of the manually segmented video segments are treated as text documents. A three-step approach that is similar to the one used in general natural language based text QA (described in the last section) is then used to identify the answers from the transcribed video segments collections. However, unlike some TREC QA systems such as Falcon (Harabagiu et al. 2000), it uses a template-based approach for question understanding and answer extraction. Such a template-based approach does not rely on complicated deep semantic analysis such as automated theorem proving in Falcon, and is believed to be more suitable for spontaneous speech text. The approach uses a parser called Conexor iSkim (Voutilainen 2000). Major verbs, nouns, noun phrases, and named entities in the question are extracted from the iSkim output. Their synonyms are found from the WordNet dictionary (Miller 1990). A query is formed using the original words, their synonyms and the named entities. The answer type of the
question is also derived and the question is filled into a question template with nine slots, including answer type, question focus, person, organization, governor, objects, number, time, and location (Figure 8) (Zhang & Nunamaker 2004).

<table>
<thead>
<tr>
<th>QUESTION TEMPLATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Answer Type (type of information a question is looking for)</td>
</tr>
<tr>
<td>Question Focus (the core noun)</td>
</tr>
<tr>
<td>Person (named person)</td>
</tr>
<tr>
<td>Organization (named organizations)</td>
</tr>
<tr>
<td>Governor (key verbs)</td>
</tr>
<tr>
<td>Objects (other noun or noun phrases)</td>
</tr>
<tr>
<td>Number (numbers)</td>
</tr>
<tr>
<td>Time (year, date, etc.)</td>
</tr>
<tr>
<td>Location (country, region, city, etc.)</td>
</tr>
</tbody>
</table>

Figure 8. Question Template from (Zhang & Nunamaker 2004)

The basic query is then sent to a Boolean information retrieval engine. The search results are processed to extract answers. Sentences in the retrieved documents are then parsed using iSkim and transformed into sentence templates (ST) in a similar form as the question template (QT). Similarity between the QT and the ST is calculated based on the combination of the following three factors (Zhang & Nunamaker 2004).

- **Matched_Slots_Score (MSS)** compares the slot values of QT with STs.

- **Same_WordSequence_Score (SWS)** computes the number of keywords in the question that appear in the same sequence in the current sentence.

- **AnswerType_Found_Score (AFS)** checks if either Answer Type or Question Focus of the question is found in the current sentence.

Finally, a sliding-window method is used to calculate the similarity between the question and the video segments. A window (a sequence of n sentences) is started from
the first sentence of the segment. The individual combined matching scores of five sentences within the current window are summed to obtain a window matching score. Then the window is moved toward the end of transcript one sentence at a time and the matching score of the current window is obtained in the same way until the last sentence of the transcript is reached. Finally the highest window matching score is selected as the relevance score of this segment to the current question. The top relevant documents are returned.

This natural language based approach works well in a limited domain such as lecture videos. In this scenario, the interpretations of questions are clear and the parsing and understanding of sentences is feasible in real time since the set of retrieved documents is small (Cao et al. 2004c). However, one disadvantage of this approach is that it completely relies on the transcribed text of the video. If transcribed using speech recognition software, errors could incur in the video transcripts, and these errors may greatly reduce the accuracy of retrieval. Therefore, this approach has to be improved to solve this problem. We borrow some idea from the video-based information retrieval research described earlier, and propose an improved video-based QA approach with the help of phonetic matching and extra knowledge sources. The proposed new approach is described in detail in the next section.
3. RESEARCH PROBLEM: VIRTUAL INTERACTION AND LEARNING WITH VIRTUAL MENTORS

The research problems to be studied in this dissertation, as mentioned earlier in the introduction chapter, are how to add interactivity into a multimedia-based e-learning system and how the effectiveness of e-learning is consequently affected by that interactivity. However, there is an important distinction between this dissertation study and the other existing interactive e-learning studies: we consider the particular context where human instructors are not available either synchronously or asynchronously. Therefore, the interaction to be studied here is actually a learner-content interaction. The innovation of this dissertation study is that we are using information technology to make this learner-content interaction simulating a learner-instructor interaction. We define this simulated learner-instructor interaction as virtual interaction, where students may feel that they interact with an instructor but they actually interact with a virtual mentor instead of a real person. A unique virtual interactivity component is designed to implement such virtual interaction and is described later in this section.

This chapter first describes an exploratory, conceptual model on how to add such virtual interaction into an e-learning system and what factors may affect the effectiveness of such an e-learning system. It then describes the architecture of a system designed based on this model. Technological innovations of this system are also introduced in this chapter. As we will see later, this system is the foundation for our investigation of the effectiveness of e-learning with virtual interactions.
3.1. Adding Interactivity into E-Learning Systems: The LVM Model

Drawing upon learning theories discussed in chapter 2 and empirical findings from previous e-learning research study (described in chapter 5), we propose an exploratory, conceptual model on how to add virtual interaction into an e-learning system and what factors may affect the effectiveness of such an e-learning system. The model is named Learning with Virtual Mentors (LVM, Figure 9). Responding to the call for a greater depth and breadth of research in the area of TML (Alavi & Leidner 2001), we propose the LVM model with a comprehensive view of the relationships among instructional strategies, information technology, students’ psychological character, environmental factors, and learning outcomes.

Three modern learning theories introduced in chapter 2, including the discovery learning theory (Bruner 1960), the constructivist theory (Bruner 1966), and the conversation theory (Pask 1975), form the theoretical foundation of the LVM model. They bring out the core concept of LVM, which is that high quality learning must incorporate a questioning and answering process, and therefore asking questions and finding answers are essential activities of virtual interaction.

However, it is noted (Jonassen 1991) that each phase of knowledge acquisition requires different types of learning and that initial knowledge acquisition is perhaps best served by classical instruction, and the constructivist learning environment is more suited to the second phase of knowledge acquisition. Therefore, as many literatures on constructivist design suggested, a mix of old and new instructional design combining both classical instruction and conversation-based constructivist learning may benefit
learners better and meet the needs of a variety of learning situations (Ertmer & Newby 1993). For example, question-based reasoning is an approach people naturally use to work through tasks and problems (Dickover 2001), and thus question-based learning can be used in just-in-time fashion after formal training is complete.

As mentioned earlier, besides 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 guides the learner in deciding how to interpret the realities, and whether to confirm, refine or revamp their mental models (Guskey 1997). Research also shows that the effects of feedback on learning varied with how the feedback is given to learners (Azevedo & Bernard 1995). "The best feedback to students is immediate, specific, and direct, and it offers explicit directions for improvement” (Guskey 1997).

The instructional strategies discussed in Chapter 2 help us identify three essential types of learning activities in a learner’s learning process in LVM: the learner receives customized instructions from the VM based on her knowledge level; she can ask questions at any time and the VM will answer her immediately; finally, at certain points in her learning process, the VM will give her a quiz to assess her knowledge and give feedback to her after the assessment is finished. The learner can decide to repeat the previous instruction, advance to the next instruction, or ask more specific questions. We believe such a cyclic process can provide learners the most flexible and natural learning and will result in comparable or even better learning outcomes as learning with real mentors.
Among the three essential learning activities, we believe that the questioning/answering (QA) process and the assessment/feedback process are the critical components of virtual interactions. Although structured instruction is very important for learning any new concept, virtual interactions especially the QA process are the most important activities for advanced learning. To implement such virtual interaction as a close simulation of real interaction, we need the help of advanced information technologies. For example, multimedia technology (mainly streaming video) is chosen to create a representation of a mentor that can let learners feel they are directly interacting with the mentor. Since the delivery of multimedia lectures and the assessment functions have been commonly studied in computer-based training systems, in this dissertation study we focus on designing and implementing the QA module that simulates the
question and answer process between real people. The technology we have selected is video-based question answering based on speech text analysis, which we discussed in Chapter 2. The improvement we made to the selected method is described later in this chapter.

Besides the instructional strategies and information technologies for implementing the virtual interaction, we also believe that there are many different factors that can influence the learning outcome when learning with virtual mentors, such as students’ psychological characters and environmental factors. As shown in Figure 9, the learning outcomes are represented by the learner’s learning performance and their perception/satisfaction to the learning process. Based on learning theories and our empirical findings in previous research (see chapter 5), we propose that there are five factors that may affect the learning outcomes: including learning phases, learning motivation/strategy, learner capabilities, system interactivity, and learning activities. We explain these constructs in detail below.

LEARNING PERFORMANCE. In most research studies about learning, a student’s learning performance is measured by his or her test score improvement. In this dissertation we give students a pretest and a posttest about the subject matter (two equal-difficulty-level tests with multiple choice questions) and measure the difference between them.

PERCEPTION/SATISFACTION. Besides the objective measure of learning performance, it is also necessary to measure students’ subjective perception/satisfaction towards the learning process. In this study we ask the students to report their perceived
learning effectiveness, their perceived system usability such as ease of use, their perceived system interactivity, their preference on a questionnaire consisting of several five-point Likert type scales self-developed (Cao et al. 2004a) or adapted from existing studies (Alavi 1994). We also propose that students’ perception/satisfaction is related to their learning performance.

LEARNING PHASES. As discussed 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 & Newby stated in (1993), 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.

LEARNING MOTIVATION/STRATEGY. A learner’s learning motivation and/or strategy can significantly affect his or her learning activities and learning achievements. “Effective learning does not occur without motivation” (Wlodkowski 1999). We expect that high-motivation learners engage in more interactive learning activities, achieve better learning performance, and report better perception/satisfaction to their learning with the system compared to those low-motivation learners.
Learning strategy or style, on the other hand, is about how learners are able to learn. For example, when a learner encounters a problem in learning, does he or she prefer to seek help or solve it independently? When he or she reviews some subject matter, does he or she like to go through everything sequentially or in different order? Although there are many instruments that have been developed to measure people’s learning strategy or learning style, we only pay attention to those that are relevant to interactive learning activities.

Three instruments are finally selected to measure the learner’s learning motivation, strategy or style. Motivated Strategies for Learning Questionnaire (MSLQ) (Pintrich et al. 1993) is used to measure the learners’ general learning motivation and strategy. It is selected because it has been validated using large sample and broadly used in many universities. It also includes modular scales that are directly related to the issue of interaction, and each scale can be used separately. Index of Learning Styles (ILS) is selected to assess learners’ preferences on three dimensions that are directly related to the issue of interaction: active/reflective, sensing/intuitive, and sequential/global (Felder 1993). Finally, Approaches to Learning and Studying Inventory (ALS) (Entwistle 1998) is chosen for measuring whether learners adopt a learning approach focused on understanding (deep learning) or a learning approach focused on reproducing (surface learning).

There are many sub-scales in each of the three instruments and we are trying to find if any of them would affect learners’ behavior (learning activities) and the resulting
learning outcome. A detailed description of each instrument and their sub-scales is given in Chapter 6.

LEARNER CAPABILITIES. We consider two aspects of learner capabilities: one is about learners’ native intellectual level; another is the learners’ prior knowledge of the subject matter. The learners’ native intellectual level can be measured by their GPA, and their prior knowledge of the subject matter can be measured by the score of a pre-test. Both of these two aspects may affect learners’ learning activities and learning outcomes. For example, we assume that higher prior knowledge will actually result in lower level of learning activities and maybe relatively low learning achievements.

SYSTEM INTERACTIVITY. This is the key construct in the LVM model. As shown in Figure 9, the classical, sequential instruction provides zero interaction in learning; the question and answering process initiated by students adds interactivity into students’ learning process; and the assessment and feedback process initiated by the Virtual Mentor (VM) 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. This construct refers to learners’ actual interactive activities in learning. Although it is difficult to measure this construct in a traditional classroom learning environment, an e-learning system can record learner’s activities in system logs conveniently. For example, the system can record how many questions a student asked during a period of time, and even capture every key stroke or mouse click that the student made. We propose that more interactive learning activities can be the result of both learner characteristics (motivation, strategy and style) and system interactivities, and they are expected to result in better learning outcomes, especially in the advanced learning phase.

To summarize, the LVM model is an exploratory, conceptual model about the relationships among e-learning system interactivity, students’ characteristics and activities, learning phases, and students’ learning outcomes. Although we assume that virtual interaction will help improve e-learning effectiveness, we want to study this problem in a broader context to see how other factors may affect this improvement of effectiveness.

3.2. System Architecture

We design a comprehensive multimedia e-learning system to implement the learning process described in the LVM model, especially the virtual interactions. The system architecture is illustrated in Figure 10. It is a three layer client/server architecture, which includes a client layer, an application layer and a database layer.
CLIENT LAYER: Students access the LVM learning environment through a Web browser. The client side is platform-independent, and only requires a web-browser, a video player and a sound card. Sample user interfaces are illustrated in Figure 11, where students would see three major components (function groups) that are implemented in the application layer. The *Virtual Classroom* comprises the major part of the user interface. It is a synchronized lecture browser which simulates the traditional classroom-learning environment by synchronizing three cells in the browser: the instructor’s video, the PowerPoint (PPT) slides, and the text transcript of the instructor’s speech.
buttons and an outline of topics are also provided so that learners can easily select any topic in the lecture at any time. The Search Tool is a text box allowing students to ask a question in everyday English or simply search by keyword (Figure 11a). The Assessment Tool provides pop-up quizzes at system specified times (Figure 11b) so that students can test their knowledge and get feedback immediately.
APPLICATION LAYER: The application layer includes an application server and a Web server, which retrieve data from the database layer and generate the presentation for all the system modules described above.

DATABASE LAYER: The Database layer includes a video streaming server and a database server. The video server stores video and audio files, and provides the video and audio streaming service. In order to provide the outline for navigation in the lecture and more importantly provide the video-based question answering, lecture raw videos have to be manually or automatically segmented into small video segments based on topic transition. This pre-step is currently conducted manually, but we are studying automatic segmentation (Lin et al. 2004) and hope to be able to do it automatically soon in the
future. In addition, speeches in lecture raw videos need to be manually or automatically transcribed into text so that video segments can be indexed based on these text transcripts. This is an important pre-step for video-based question answering that we will describe in detail in the next section (Zhang & Nunamaker 2004). As depicted in Figure 10, the meta-data about video segments (e.g., start time and end time in the raw video) and lectures, transcripts, topics, examples, quizzes, and PPT slides are all stored and organized in the database server.

Since the virtual classroom and the assessment functions have been commonly studied and implemented in computer-based training systems, they are not described in detail in this dissertation study. Instead we focus on designing and implementing the QA module. The major technology innovation here is the improvement of the video-based question answering technology (Zhang & Nunamaker 2004) using extra knowledge sources. By using proper natural language processing technologies (e.g. robust parsing, phonetic correction, and semantic analysis) to pre-process and index the lecture video transcripts, as well as their associated slides or lecture notes (as extra knowledge sources), the QA module will allow students to type in a natural language question and receive a segment of the video playing back as the answer.

3.2.1. Using PowerPoint Slides to Improve Video-based Question Answering

Specifically, in the QA module, a natural language question submitted by a user is analyzed to derive a query. Segments that may answer the question will then be extracted from multiple video sources by analyzing the text transcripts using natural language understanding. The segments are ranked by their relevancy to the question and the most
relevant answer is directly played back to the user. The process has three major steps and they are illustrated in Figure 12. The development of the QA module is based on the assumption that video segmentation has been done automatically or manually.

1) VIDEO PREPROCESSING: In this step, the text transcripts of lecture videos are generated by a speaker-independent speech recognition tool: Virage VideoLogger® (http://www.virage.com). This tool also generates time stamps that synchronize the video stream with the transcribed text at word level. However, speech recognition errors could occur in the video transcripts, and these errors may greatly reduce the accuracy of retrieval. We solve this problem by doing transcript correction based on phonetic matching, a method described in (Yang et al. 2003). A list of domain concept words is developed based on content in the domain knowledge base. Words in transcripts are converted to phonetic sounds and are compared to the phonetic sequence of the words on
this list. Similar sound words in the transcripts will be changed to the one on the word list. Finally, the corrected video transcripts are indexed and stored in a database.

2) QUESTION PROCESSING: We use a parser called Connexor iSkim (Voutilainen 2000) to parse an input natural language question. Major verbs, nouns and noun phrases in the question are extracted from the iSkim output; their synonyms are found from the WordNet dictionary (Miller 1990); and a query is formed using the original words, their synonyms and the named entities. Besides this basic query, we also derive the answer type of the question (e.g. the answer type for “Who is” question is “person”), and fill the question into a question template with seven slots, including answer type, governor, object, person, time, and so on (Zhang & Nunamaker 2004). The basic query and the question template (QT) are sent to the next step.

3) ANSWER EXTRACTION: In this step, the basic query is sent to a Boolean information retrieval engine. The search results, a list of relevant transcripts, are processed to extract answers. Sentences in these relevant transcripts are parsed using iSkim and transformed into sentence templates (ST) in a similar form as the question template. Relevancy between the QT and the ST is calculated. Relevancy between the basic query and domain knowledge sources is also calculated and combined with the former. Specifically, we use text extracted from PowerPoint slides as a resource of domain knowledge to improve the answer extraction performance (Cao & Nunamaker 2004). Shown in Figure 11, a PowerPoint slide is usually associated with a lecture video for a period of time. Obviously, for any video segment in this period, if both its transcript
and the text in its associated slide are relevant to a question, this video segment is more relevant to the question than other video segments with only transcript or relevant slides. 

Therefore, the total relevancy score between a video segment Segij (jth segment in ith video) and the question is calculated using the following formula:

\[
\text{Seg}_{ij} = \alpha \sum_{k=1}^{s} W_k + \beta \sum_{s} W_s
\]

where \( \alpha + \beta = 1 \), \( W_k \) is the relevancy score between the question and the kth sentence in the transcript segment, and \( W_s \) is the relevancy score between the question and the text of the sth slide associated with the video segment.

Finally, a sliding-window method is used to calculate the total relevancy between the question and each video segment in the transcript. The top five segments are returned as answers to the user and the first answer, which is assumed to be the most relevant answer, is played back automatically. This design closely simulates the scenario of questioning and answering process between real people, because the students could ask a question and watch the virtual mentor directly telling him or her the answer. However, this presentation method is also risky to some extent. If the answer extraction is not accurate enough, students will be given a wrong answer in the beginning and they have to select the other answers manually from the answer list. This may thus affect the students’ perception of the system effectiveness. Therefore, it is important to know how good our QA algorithm is before we can use this auto-playback presentation method.

3.2.1.1 Evaluation of Answer Extraction

The QA module has been evaluated using a professional training course (Deception Detection) as the test bed.
To evaluate answer extraction, a group of PhD students who are experts in the respective domain were invited to act as assessors. Each assessor created a set of questions and submitted the questions to the answer extraction module. A ranked list of answer segments plus the link to the original video documents were returned to the assessors. The assessors then watched the top 5 segments returned in the ranked list and made a binary decision as to whether each segment was relevant to the question. Finally, a mean reciprocal answer rank (MRR) was calculated to evaluate the QA module according to the following formula (Voorhees 1999).

\[
MRR = \frac{1}{\sum_{i=1}^{\#\text{questions}} \frac{1}{\text{answer}_i\text{rank}}}
\]

Where \( \text{answer}_i\text{rank} \) is the rank of the first correct answer for question \( i \), and if the answer is found at multiple ranks, the best rank will be used. If no relevant answer is found in the top 5, the score for that particular question is zero. The highest MRR score is 1 and higher scores denote more system effectiveness. MRR scores of 0.5 can be roughly interpreted as either “in average” the correct answer being the second answer found by the system, or half of the questions getting the first answer correct. We have compared the results using three different sets of transcripts, and for each set, we compared the method of using PPT slides in answer extraction to the one without using PPT slides. Our results are listed in Table 1. We conclude that our QA algorithm worked best for human generated transcripts plus the help from PPT slides (MRR = 0.561). Considering that the existing reported evaluation results for some of the best text-based QA systems are merely above 0.6 (Voorhees 1999), we think our result was reasonably good. In addition, QA based on the transcripts directly generated by the speech recognition software...
PPT slides had even higher MRR than QA based on only human generated transcripts. This is encouraging and shows that it is possible to get satisfying retrieval results without the time-consuming human correction by just using PPT slides. However, it is surprising that although the MRR for QA on transcripts with automatic correction was higher than QA on transcripts without error correction, using PPT slides did not increase the MRR as much as on the transcripts without error correction. One possible explanation is that we used the same set of PPT slides in transcript correction so there was not much extra knowledge in answer extraction to boost the performance. Also, the transcripts might be over-corrected and new errors might be introduced. More test questions and larger test collections may be needed in the future to derive more reliable conclusions.

<table>
<thead>
<tr>
<th>Transcript</th>
<th>MRR</th>
<th>Transcript + PPT Slides</th>
</tr>
</thead>
<tbody>
<tr>
<td>without error correction</td>
<td>0.423</td>
<td>0.524</td>
</tr>
<tr>
<td>with error correction</td>
<td>0.476</td>
<td>0.517</td>
</tr>
<tr>
<td>human generated</td>
<td>0.511</td>
<td>0.561</td>
</tr>
</tbody>
</table>
4. A MULTI-METHODOLOGICAL RESEARCH APPROACH

This dissertation research takes a multi-methodological research approach. Such a synergistic and multi-methodological research approach has been requested and debated for decades in the Information Systems (IS) research field. Although many empirical research methods traditionally used in social science have been applied to the IS field, such as case study, experiment, survey, action research and others (Benbasat 1984; Galliers & Land 1987), research shows that more than 80% of individual IS research studies in the 1990’s only used a single research method (Nandhakumar & Jones 1997; Walsham 1995). This single-methodological approach, although argued by some researchers as more rigorous (Mingers 2001), usually results in only a limited view of a particular research situation. In addition, those commonly used, social science based methodologies are not appropriate in studying the core subject matter of IS - the information technology (IT) artifact. After several years of a general shift in IS research away from technological to managerial and organizational issues, IS researchers began to call for a return to an exploration of the unavoidable technical component of all IS research (Orlikowski & Iacono 2001). Therefore, we argue that a research methodology for studying the IT artifact is critical for IS research, and combining different research methodologies may yield convergent validation and a richer understanding of the phenomenon under investigation. The two Agent99 Trainer studies presented in the next chapter and the latest LVM study described in chapter 6 clearly demonstrated the benefits of such a multi-methodological research approach.
A number of multi-methodological IS research frameworks or guidelines have been proposed to guide the combination of multiple research methods in IS research. Nunamaker, Chen, and Purdin proposed a framework of multi-methodological IS research approaches (Nunamaker et al. 1991). They recommended the system development methodology as a complement to the traditional empirical IS research methodologies. Mingers (Mingers 2001) proposed some practical guidance for choosing and combining different empirical IS research methods. Recently, Hevner, March, Park, and Ram (Hevner & March 2003; Hevner et al. 2004) further developed this multi-methodological view of IS research into a multi-paradigm framework. They classified IS research into two paradigms: behavioral-science researchers focus on developing and justifying theories that explain the impacts of the technologies to business, while design-science researchers focus on building artifacts to meet business needs and evaluating the utilities of the artifacts (Hevner & March 2003). They argued that better IS research should use a synergistic approach combining both the paradigms.

Nunamaker et al. (1991) advocate that IS research could be better conducted with a multi-methodological approach that includes the complementary strategies of: theory building, systems development, observation and experimentation (Figure 13). Although the legitimacy of systems development as a valid IS research methodology has been debated extensively, they justify systems development as a research methodology that fits comfortably into the category of applied science, belonging to the engineering, developmental and formulative types of research. With the driving force and goal of IS research – theory building in the center, the IS research can be viewed as an iterative
process of studying the IT artifact itself (system development) and studying the impact of the IT artifact from different perspectives (experimentation and observation). The triangulation of the results from different methodologies has the potential to lead to more powerful and insightful findings.

Figure 13. Multi-methodological Approach to IS research (Nunamaker 1992)

Hevner and March (2003) propose functionally the same framework for IS research in a different way (Figure 14). They define the environment as a problem space in which the phenomenon of interest resides, and the knowledge base as a composition of IS research foundations and methodologies. The center of this framework is the IS research that applies appropriate foundations and methodologies from the knowledge base, according to business needs from the environment. They argue that better IS research can be achieved through the combination and integration of two paradigms: behavioral-science
and design-science. Such integrated IS research is conducted in two phases: develop theories/ build artifacts, and justify theories/evaluate artifacts. By using multiple research methods, the justifying/evaluating activities can result in better refinement and reassessment of the theory or artifact.

This dissertation research mainly follows the framework described in Figure 13 (Nunamaker 1992). One major benefit of such a multi-methodological IS research approach blending both the design-science and behavior-science paradigms is that the results from multiple research methods can complement each other, resulting in better
theories. The system development methodology (or design science effort), particularly, contributes to better theories because 1) it can be an experimental proof of concept and 2) it can expose relationships between the elements of the IT artifact which enable certain behaviors and constrain others (Purao 2002).

In addition, based on our experiences from the Agent99 Trainer studies (details presented in Chapter 5), we discovered that there can be two different types of interactions between different research methods and both interactions can lead to better IS research. These two types of interactions are illustrated in Figure 15, and the resultant new vision of the benefits of multi-methodological IS research is explained below.

As this figure illustrates, research methods used for different research goals (testing theory/truth or evaluating system utility) can interact with each other in two ways. In Interaction 1, the Result-Result (R-R) interaction, the results from one method can help the interpretation of results from another method, and vice versa. The R-R interaction
usually results in either better understanding of the theory, better improvement of the system design or both. For example, results from a system usability test may help researchers understand the system limitations and thus help them interpret the non-significant results from a theory-testing experiment. The R-R interaction is the commonly recognized benefit of a multi-methodological research.

Interaction 2, the Result Design (R-D) interaction, however, results from one method in one paradigm interacting with the research design of another method in another paradigm. The R-D interaction can result in better research design. Since the appropriate research method and research design is the key for successful research, we view the R-D interaction to be even more important than the R-R interaction. For example, results from a system usability test can help researchers discover any unintended variation of the actual use of the system, and thus help them identify some hidden problems of the design of a theory-testing experiment. Without this R-D interaction, the same experiment might be conducted repeatedly without finding the real truth, because the users do not actually use the system in the way that the researchers expected. Therefore, the R-D interaction between multiple methods is critical to IS research.

We also emphasize that it is better that the interactions are between methods from the two different research paradigms. As Hevner and his colleagues stated, “…An artifact may have utility because of some yet undiscovered truth. A theory may not yet be developed to the point where its truth can be incorporated into design…” (p. 7, (Hevner et al. 2004)). For researchers doing the behavioral-science type of IS research using a given IS system, it is better that they also choose a method to evaluate the system utility.
The results from this extra study could provide deep understanding of how the information system is used in the experiment, and thus give them great power in validating their major research design and results. On the other hand, for researchers doing design-science type of IS research, besides building a system according to certain business requirements and evaluating its utility, they might also want to choose a method to do theory testing to find which factor or principle accounts for the utility of the system. Again, results from this complementary method could let them fine-tune their system utility evaluation methods. Overall, such a multi-paradigm, multi-methodological IS research framework will result in better IS theories and/or better information systems.

As a result, we conduct this dissertation research with a multi-methodological approach because of the great benefits of triangulation of research methods. Taking the LVM model proposed in Chapter 3 as the center of this research (theory), we start by creating an IT artifact through system development, and then conduct positivist experimentations and interpretive observations. The results from experimentations and observations are interpreted and triangulated, becoming the basis for new theorizing and a new cycle begins. The three major activities in this iterative process are described in detail below.

4.1. System Development

System development always involves creation of some artifacts “…which may be either purely conceptual artifacts (models, frameworks and procedures) or more technical artifacts with a ‘physical’ realization (e.g., software)” ((Iivari et al. 1998), p.175). In the context of IS research, Nunamaker et al. argue that the systems development
methodology consists of five parts: 1) constructing a conceptual framework, 2) constructing the architecture of the system, 3) analyzing and designing the system, 4) prototyping, and 5) evaluating the system (Nunamaker et al. 1991). The conceptual framework is usually derived from the theory base of the entire study, consisting of the system functionalities and requirements. The system architecture defines the functionalities of system components and interrelationships among them. The analyzing, designing and prototyping process determines the specific system implementation, for example the specific design and implementation of algorithms, databases and user interfaces. Implementation of a system (prototyping) can be used to demonstrate the feasibility of the conceptual model. Findings from empirical evaluations of the functionality and the usability of a system can be used to redesign the system, or finally modify the concepts and theories from which the systems are derived.

**Figure 16. A General Design Research Methodology**
A very similar process of design research methodology built upon the work of Nunamaker et al. is described in (Vaishnavi & Kuechler 2004). As illustrated in Figure 16, a typical design research effort proceeds with 1) Awareness of Problem 2) Suggestion of a Tentative Design 3) Development 4) Evaluation and 5) Conclusion of the effort and the knowledge gained in the effort. This process can be easily mapped into the one described in (Nunamaker et al. 1991). Again, the most important part of a system development or design research process is the analysis, design, and implementation of a system.

4.2. Experimentation

From the behavioral-science perspective, IS research is viewed to be a branch of the social sciences (Hevner & March 2003). Growing from natural science research methods, the primary goal of IS research in this paradigm is to evaluate the truth of hypotheses and to validate theories developed to inform the design of information systems. Evaluation and validation activities or research methods in the behavioral-science paradigm include case studies, experiments, and field studies. These activities enable researchers to identify strengths and weaknesses in their theories, as well as often providing new insights.

Among the many research methods, experimentation is the predominant one in behavioral sciences. An experiment is a controlled experience of participation in a designed event (Black 2002). As such, the experiment has many characteristics that distinguish it from other evaluation methods. A classical experiment involves a representative sample of participants, randomly assigned to multiple groups which are given different treatments (Black 2002). The treatment groups are then measured and
compared on the outcome variable(s) of interest to evaluate theory. McGrath (1982) suggests that a well-designed experiment can serve three underlying research goals: 1) control of extraneous influences, 2) control (i.e., manipulation) of independent variable(s), and 3) control of the setting (McGrath et al. 1982). This significant degree of control inherent in experiments allows the discovery of causal relationships, the critical component of theories used to explain and predict phenomena. Therefore, experimentation is a common tool for theory testing.

Although there are innumerable permutations of experiments, McGrath identifies three broad categories: laboratory experiments, experimental simulations, and field experiments (McGrath et al. 1982). These categories differ in their level of precision of measurement of behavior (which includes control of external influences), realism of the context, and generalizability of the findings.

Laboratory experiments are conducted in deliberately contrived settings that are not existentially real for the participants. Researchers attempt to create a universal setting for the behaviors under study in the laboratory. In experimental simulations, an attempt is made to create a realism of content in the lab, an unreal context. The research advantage of both is maximizing the precision of the measurement of behavior (independent variable) (McGrath, 1982). In contrast, field experiments attempt to impose the controls of a laboratory experiment in real settings (i.e., in the field). As a consequence, field experiments trade some level of precision in measurement for realism of the context and resultant participant behaviors.
For all types of experiments, (Campbell & Stanley 1963) identify three true experimental designs: 1) the pre-test-post-test control group design, 2) the Solomon four-group design, and 3) the post-test-only control group design. The key characteristics of all of these designs are random assignment to control external variance and multiple treatment groups (a control group is simply a void treatment group) to provide a source of comparison. These designs serve as the basis for most, if not all, experimental research in behavioral-science, as well as quasi-experimental research. Experimental designs are guided by theories and facilitated by system development. Results from experimentation may be used to refine theories and improve systems.

4.3. Observation

Observation includes research methodologies such as case studies, field studies, and sample surveys that are unobtrusive research operations (Nunamaker et al. 1991). Unlike experimentation, observations can be conducted in more natural research settings, without manipulation of constructs. The focus of observation is not to test causal relationships among constructs, but to “get a general feeling for what is involved” in a research domain (Nunamaker et al. 1991). As stated earlier, the results of observations may not only help researchers explain results from experimentation, but also help them refine the experimental design. The observations can also facilitate system development.

One major use of observation in IS research is to test system usability. From design-science perspective, the crucial goal of IS research is to evaluate the quality of the artifact – the information systems. Such an evaluation requires the definition of appropriate metrics or attributes of quality. Some researchers view usability as one of the quality
attributes, and list functionality, completeness, consistency, accuracy, reliability, performance, and others as quality attributes (Hevner et al. 2004). However, here we take a broader but commonly accepted view of usability. Following the definition by Dumas and Redish (Dumas & Redish 1993), usability measures whether people who use a system can do so quickly and easily to accomplish their own tasks. Therefore, usability can work as an integrated measure of IS quality. Attributes like functionality, accuracy, performance, reliability and others can be viewed as usability attributes, as suggested in (Nielsen 1993).

The usability of an information system can be evaluated on a set of usability attribute measures described above. The overall goal of the evaluation, however, is to not only assess the system usability but also provide guidelines for improving the usability and therefore the entire system quality. Nielsen (1993) described several common methods for evaluating usability, such as heuristic evaluation, thinking aloud, questionnaires and interviews. Although experimentation can also be used for usability testing, most of the activities or methods of usability evaluation are considered as “observation”, because they require the researchers to carefully observe the users’ experiences in using the system. This observation of the actual use of the system can be either from an objective perspective, such as system log or thinking aloud; or from a subjective perspective, such as users’ self-reports in a questionnaire.

Among the common usability testing methods, questionnaires/surveys and interviews are the two most popular methods for investigating users’ subjective preferences and attitudes (Nielsen 1993). Both of them are easy to implement, as well as very useful for
investigating which system features users particularly like or dislike, and are commonly used in research studies to evaluate a system’s perceived usability. Questionnaires require users to answer a fixed set of questions about the system. In contrast, interviews are conducted by an interviewer who asks the users questions about the system and records the users’ answers. Interviews are usually considered to have more flexibility than questionnaires, but conducting interviews is also much more labor intensive. Most objective usability testing methods also suffer from this problem. Since users can complete questionnaires independently, questionnaires save time in comparison to interviews and can be administrated to several users simultaneously, unconstrained by the need for an interviewer or specific labs and/or equipments. Therefore, although usability can be evaluated in many different ways, usability questionnaires are chosen by most researchers when conducting large scale studies. Therefore, it is easier to couple a usability testing, if using questionnaires, with large scale experiments.

In this dissertation study, we develop the system iteratively by designing common system architecture and implementing different prototypes based on specific conditions and feedbacks. We test our conceptual model using a field experiment with a pre-test-post-test control group design. We also observe users’ behaviors in using the system as well as the system usability with a survey questionnaire and a semi-structured interview. The different research methodologies triangulate with each other and provide us the best insight into the phenomena to be studied.
5. EARLY EXPLORATION OF THE LVM MODEL: THE AGENT99 STUDY

The development of the LVM conceptual model went through a long period. As stated earlier, the LVM model is the result of not only learning theories but also empirical findings. The empirical study that helped us form the LVM model is the Agent99 Trainer (A99) study, which is a study about using a computer-based training system to train people to detect deception (George et al. 2004b).

The A99 study has two phases. In the first phase, a simple multimedia learning system with limited interactive functionalities was built and used for deception detection (DD) training. The findings from this phase, combined with our review of the learning theories, prompted our propositions about the virtual interactions in the LVM model. In the second phase, the effect of virtual interactions was explored but with no success. The findings, especially the observations in this phase of study revealed another important factor in the LVM model – the leaning phases. The two phases of the A99 study allowed us to see the deficiency of individual factors in explaining the effectiveness of e-learning systems, and finally prompted our comprehensive view of this problem – the LVM model.

The A99 study was initiated with a specific task: training people to detect deception. Deception is a complex human behavior. The high failure rate of DD by humans, as well as the shortage of deception detection experts (available trainers), render the task of training people to detect deception extremely challenging (Frank & Feeley 2002; Miller & Stiff 1993). Therefore, there were two major goals of this study: 1) to justify that training using reliable cues or indicators of deception can improve humans’ capability of
detecting deceptions; and 2) to build a computer-based training system to implement such DD training programs, as a more efficient and cost effective alternative of traditional classroom training.

To accomplish these two goals, we conducted the study based on Nunamaker et al.’s multi-methodological IS research framework (Nunamaker et al. 1991). Particularly, we iteratively developed system prototypes, evaluated the system usability using questionnaires, and tested the theoretical hypotheses using field experiments. As introduced earlier, there were two phases in this study and they were described in detail below.

5.1. Agent99 Phase I: Build A Virtual Classroom For Deception Detection

The A99 study phase I is our first trial of building a computer-based training system for deception detection. The LVM model was not proposed at that time. So we simply started with the review of deception training literature and the training system development in 2001. The first version of the training program and the prototype of the system (named Agent99 Trainer) were evaluated in 2002. Field experiments were done to test our theoretical hypotheses and usability surveys were used to test the utility of this prototype (Cao et al. 2003; Cao et al. 2004a; George et al. 2004b). However, the results from the 2002 experiments did not support all of our hypotheses. Fortunately, several limitations of the first system were exposed from the usability testing and suggested that the failure of hypotheses testing might be caused by the limitation of system functionalities, especially the lack of interactions. Therefore, several improvements were
made to the system after the 2002 evaluation and we started the A99 study phase II, with one goal of exploring the effect of interactive functionalities in the A99 system.

In the following sections, we describe the system development and the evaluation studies conducted in the phase I.

5.1.1. System Development

Our review of the DD training literature (Cao et al. 2003; Frank & Feeley 2002) identified three critical components of DD training: explicit instruction, practice, and feedback. Since DD is complex, a deep understanding of cues requires extensive experience and high levels of cognitive processing. Thus, not only is it important to incorporate the principles of a good training system, but also make the training learner-centered by providing navigational facilities to different modules, library of examples with analysis, and assessment examples with immediate feedback and analysis. With these goals in mind, we developed our first prototype of A99, a computer-based training system for DD. This prototype implemented the three critical components of DD training in two modules (Lin et al. 2003b): Watch Lecture and View Example with Analysis. The system architecture was similar to the one shown in Chapter 3 (Figure 10), except that there were no QA and assessment modules. All modules in this prototype were implemented using HTML, Java servlet, Java script and VB script.

The Watch Lecture module (Figure 17) provided explicit instruction on deception cues through a combination of synchronized media including the expert video lecture, presentation slides, and lecture notes. The instruction specifically focused on five behavioral categories of cues: arousal, emotion, cognitive effort, memory process, and
communication tactics. The lecture was segmented into different topics. Navigational buttons and pull-down menus allowed users to maneuver to any topic or example. The View Example with Analysis module (Figure 18) provided real life examples and scenarios on judging the veracity of communications based on the lectures. It also provided immediate feedback and analysis of the deception cues illustrated.

Figure 17. Agent99 Trainer I Watch Lecture
5.1.2. Field Experiment for Theory Testing

5.1.2.1 Experiment Design

Training for DD has been studied for years; however, results of past training are mixed (Frank & Feeley 2002). Many training experiments failed to produce appreciable improvements in humans’ ability of detecting deception. Therefore, the first question to be answered in our research study was: can training using reliable cues or indicators of deception truly help people better detect deception? On the other hand, previous research showed that computer-based training, especially web-based training can provide benefits such as self-paced, repeatable, and cost-effective access to the curriculum over the traditional classroom training (Lin et al. 2003a). Although these benefits were
demonstrated in certain general training or learning domains, the question that remains unanswered was whether they will still be true in the DD training domain? Therefore, the two hypotheses we tested in the A99 study phase I were as follows: 1) learners receiving DD training would improve their deception detection accuracy more than if they had received no training; and 2) learners receiving training through A99 would improve their DD capability more than if they had received training by an instructor-led, lecture-based training.

Field experimentation was chosen as our research method to test these two hypotheses, mainly because its power of discovering causal relationships that we mentioned in section 2. The study was conducted in fall 2002 on a large U.S. Air Force (USAF) facility located in the United States. A total of 125 officers participated as subjects and the DD training was part of their training program. In this experiment, rather than a single lecture, a series of three training sessions was offered: 1) an introduction to DD (Intro), 2) cues used to detect deception (Cues), and 3) heuristics for decision making that are susceptible to deception (Heuristics). Each session was designed to last for one hour. The subjects were assigned to “blocks,” or classes, made up of sixteen officers, that were randomly assigned to either the control or to one of three treatment groups. The control group received no training, but they completed the same measurement instruments as the experimental subjects. The treatments were different in the second training session. For one treatment, the lecture-only training groups, traditional lectures were used. For the second treatment, the computer-based training groups, the A99 system was used exclusively in the second session. For the third treatment, the
“combination” group, a combination of lecture and A99 was used. All lectures in all treatments were supported with the same PowerPoint presentations and examples.

Table 2. Experimental Design of Experiment 1

<table>
<thead>
<tr>
<th>Training Sessions</th>
<th>Treatment Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intro</td>
<td>No Training</td>
</tr>
<tr>
<td></td>
<td>Lecture Only</td>
</tr>
<tr>
<td>Cues</td>
<td>No Training</td>
</tr>
<tr>
<td></td>
<td>Lecture Only</td>
</tr>
<tr>
<td></td>
<td>A99 Only</td>
</tr>
<tr>
<td></td>
<td>Combination</td>
</tr>
<tr>
<td></td>
<td>Pre-tests</td>
</tr>
<tr>
<td></td>
<td>Pre-tests</td>
</tr>
<tr>
<td></td>
<td>Pre-tests</td>
</tr>
<tr>
<td></td>
<td>Pre-tests</td>
</tr>
<tr>
<td>1 Hour Training</td>
<td>(for the control group with No Training, there was a 1 hour break)</td>
</tr>
<tr>
<td></td>
<td>Post-tests</td>
</tr>
<tr>
<td></td>
<td>Post-tests</td>
</tr>
<tr>
<td></td>
<td>Post-tests</td>
</tr>
<tr>
<td></td>
<td>Post-tests</td>
</tr>
<tr>
<td>Heuristics</td>
<td>No Training</td>
</tr>
<tr>
<td></td>
<td>Lecture Only</td>
</tr>
<tr>
<td></td>
<td>System Try Out</td>
</tr>
<tr>
<td></td>
<td>Usability Questionnaire</td>
</tr>
</tbody>
</table>

The basic procedures for the training sessions began with subjects completion of pre-test instruments, including a knowledge pre-test (multiple-choice questions about the concepts of deception) and a DD accuracy pre-test (judging the veracity of six recorded interview examples), both directed at determining how much of the subject matter the trainees understood entering the session. Subjects were then trained for approximately one hour, except for control subjects, who were given a one hour break in the interim. Afterwards, all subjects completed a knowledge post-test, made up of the same questions as the pre-test but in a different order, and a DD accuracy post-test, similar to the pre-test but consisting of different examples. The experimental design is summarized in Table 2.

To “debug” the A99 and validate the experimental procedure, two pilot studies were conducted prior to the 2002 field experiment. Participants were undergraduate students enrolled in summer classes offered by the Management Information Systems department at a research 1 university in the Southwest. The procedure followed was the same as in
the field experiment, except that the experiment design was simplified. Only one training
session – the cues lecture – was tested, and only the DD accuracy tests were used as pre-
tests and post-tests, but the examples in these tests were different from those used in the
pre-tests and post-tests of the experiment. There were three instead of four treatment
groups in Pilot 1: control, A99, and traditional lecture. Since the results of the first pilot
study demonstrated that the pre-test and post-test were statistically equivalent and that
there was no practice effect between the pre-test and post-test, only two treatment groups:
A99 and traditional lecture, were tested in Pilot 2.

5.1.2.2 Results

Two dependent variables were measured in the field experiment: the DD accuracy
measured by the number of correct judgments divided by the total number of test cases in
each DD accuracy test; and the DD knowledge measured by the number of correct
answers divided by the total number of multiple choice questions in each knowledge test.
Independent t-tests showed that DD knowledge improvements in the treatment groups
differed from the control group for all three sessions (Intro: t(113)=-8.921, p < .001; Cues:
t(113)=-4.54, p < .001; Heuristics: t(113)=-7.536, p < .001); revealed that the subjects’
DD knowledge was significantly improved in the three training session groups, while the
control group did not improve, just as we predicted. However, there were no significant
differences between A99 and the other delivery modes on DD knowledge tests. In
addition, there was no significant improvement on DD accuracy in all groups, which did
not support our prediction (detailed data analysis can be found in (George et al. 2004b)).
Therefore, a big question for us at that time was: why did the results not support all our predictions? Was it the case that training could not improve DD accuracy? Was it the case that computer-based training could not provide extra benefits to traditional classroom learning? Or was it just because some problems of our design prohibited us from finding the real truth? If we had only conducted the field experiment, we would never know the answers and we might have to repeat the experiment again and again. Fortunately, both the pilot studies and the usability testing integrated in the experiment helped us in explaining these results. The results from pilot 2 showed a significant improvement on \textit{DD accuracy} (A99: Mpretest = .4222, Mposttest = .6889; Lecture: Mpretest = .4405, Mposttest = .6429; F(1, 27) = 32.29, p < 0.01; detailed data analysis can be found in (Cao et al. 2003)). Since the experiment procedure and the training contents were almost the same for the field experiment and the pilot studies, and only the examples used in the DD accuracy pre-tests and post-tests were different, the non-improvement of detection accuracy could be a result of inappropriate pre-test and post-test questions. The usability testing, on the other hand, helped us understand why there was no difference between the A99 and the traditional classroom training.

5.1.3. Usability Testing

A usability test was conducted to assess whether the design of the first prototype of the A99 met its objectives. Since users’ subjective attitudes toward the system were the major interest of this study, attributes such as learnability, ease of use, satisfaction, effectiveness/usefulness were selected to be measured from several different perspectives (Cao et al. 2004a). A questionnaire was employed to investigate these attitudes. The
questionnaire was iteratively revised after each pilot study or experiment, but all versions contained both closed-ended items rated on a 7-point Likert scale and open-ended questions. The closed-ended items were either directly from an existing validated measure called the System Usability Scale (Brooke 1996); or adapted from other existing validated measures such as Questionnaire for User Interface Satisfaction (Chin et al. 1988) and Perceived Usefulness and Ease of Use (Davis 1989); or developed by us regarding some specific system features.

The questionnaires were administrated at the end of the pilot studies or the experiment. For the field experiment, since only two treatment groups used A99 during session 2 in the experiment, the other two groups (control and lecture) were given 30 minutes to use A99 before completing the questionnaire (see Table 2). All groups completed the questionnaire online after all other phases of the experiment were completed. Unfortunately, because of a computer programming error, the online responses to the closed-ended questions were miscoded and non-interpretable. Therefore, only the responses to the open-ended questions in the field experiment were analyzed. For the pilot studies, participants in the A99 group completed the questionnaire and the responses for both closed-ended and open-ended questions were analyzed.

5.1.3.1 Results

For both Pilot 1 and Pilot 2 data, means were calculated on each closed-ended item and on a smaller number of composite measures extracted by a principle component factor analysis that represented the major usability attributes. The results from both Pilot 1 and 2 were highly positive with all means less than or close to 3 (“Completely Agree” –
“Slightly Agree”) (see (Cao et al. 2004a) for more details). This indicated that most of the students agreed that A99 had good learnability, was easy to use, and so on. Both the delivery method (structured and synchronized multimedia) and the training contents (especially examples and analyses) used in A99 were perceived as very effective for DD training. One of the major features of the A99 design – the self-paced learning method – was also evaluated very favorably. However, despite the positive ratings, subjects using A99 for training did not perform better than those receiving traditional training. Why not?

The qualitative data collected from the open-ended items in the questionnaires provided important information about the users’ actual usage of the system in the experiments, and thus helped answer the above question. Responses on the open-ended questions showed that users liked that the system was easy to use, provided a structured and synchronized multimedia lecture, and provided user control of their learning pace. Furthermore, responses indicated that the examples and analyses were very important to the DD training effectiveness. However, they also described some problems with the system such as inappropriate navigation; the relatively low audio/video quality caused by either the file compression necessary for Web-based delivery or non-professional video production; and most importantly the lack of interaction between users and the instructor in the system. These negative comments indicated that perhaps the Agent99 group did not out-perform the classroom training group in our experiments, because of system weaknesses, especially the lack of interaction.
5.2. Agent99 Phase 2: Exploring the Effect of Virtual Interaction

Based on all the results and lessons we learned from the 2002 experiment and usability testing, we not only redesigned the A99 system, but also redesigned the experiment and refined the usability test. The improved system and the hypotheses were evaluated in 2003, again, using a combination of different research methods including field experiments and usability surveys (Cao et al. 2004b; George et al. 2004a). The results from the 2003 usability studies not only revealed additional system limitations, but also helped us discover some weaknesses of the experiment design that need to be redesigned to better test the theoretical hypotheses. In addition, the observations from this phase of the A99 study helped us discovered another factor that might affect the effectiveness of the learning system – learning phases, and therefore helped us finally view this problem in a more comprehensive way and propose the LVM model. These iterative interactions between different research methods again demonstrated the benefits of a multi-methodological approach to IS research.

5.2.1. System Development

The second prototype of A99 was designed to overcome the shortcomings of the first prototype, which were revealed with the prior experiments and the usability studies. To deal with the bandwidth restriction problem and the fact that the intended users of this system, Air Force officers, posed problems regarding networking security concerns from the military, the delivery media of the second prototype was changed from Web to CD-ROMs. The system architecture was then changed from client-server architecture to Microsoft Windows-based stand-alone applications. However the modules were the same
as the ones shown in the conceptual system architecture described in Chapter 3 (Figure 10). The Watch Lecture and View Example with Analysis modules in the prototype I were integrated into a single Virtual Classroom module with more fine-tuned synchronization between the lecture video, presentation slides, and lecture notes (Figure 19). The interface was revamped based on the user comments from the study phase I. Examples were now embedded in the lecture video, while the pull-down menu was replaced by an outline-type navigation menu facilitating better maneuverability to choose any topic. Most importantly, two new modules called Search Tools (Figure 19) and Assessment Tool (Figure 20) were designed to provide better interaction and assessment capabilities. The Search Tool included components such as Keyword Search and Ask A Question (AAQ) to facilitate user-driven information delivery. The search query resulted in a list of relevant video segments in the lecture video. Students might select any segment from this list to watch the answer. The Assessment Tool included popup quizzes at certain break points to facilitate user evaluation and immediate feedback through correct answers. Thus the new design took into consideration the shortcomings identified by the study phase I and improved the earlier prototype. All modules in this prototype were re-coded using Visual Basic.
5.2.2. Field Experiment for Theory Testing

Because the results from the study phase I did not fully support our hypotheses, especially the hypothesis that the A99 can provide better learning achievement than the
traditional training, a field experiment was conducted again at the same USAF facility in 2003. However, the design was changed according to the findings from the study phase I and the subsequent system changes. First, for the classroom training treatment, video lectures were used instead of live instructors to standardize the presentation order and content. The video production was conducted in a more professional way to improve the audio/video quality. Second, to test whether the previous year’s failure was caused by the partial implementation of A99 and whether the newly added interactive functionalities could improve the training effectiveness, the other treatment groups were designed to use three different versions of the A99. Finally, although both the literature and the results from the 2002 usability testing indicated that examples and analyses were very important for making DD training effective, we wanted to know explicitly whether more examples would make the training more effective. Therefore the participants, 180 Air Force officers, were randomly assigned to the following 5 treatment groups, each more sophisticated than the last. Our expectation was that participant performance should be increasingly better as one moved from treatment 1 to treatment 5 (George et al. 2004a).

Treatment 1: Video Only. Users watched a video of an expert talking about deception, with PowerPoint slides and examples cut in the video. The presentation order in this video is pre-determined by the instructor.

Treatment 2: Linear A99. Lecture outline, expert video, PowerPoint slides and lecture notes were shown on the A99 interface. PowerPoint slides and lecture notes were synchronized with the lecture video. However, users could not click on the lecture
outline to change topics. They had to follow the pre-determined instruction pace and
sequence.

_Treatment 3: A99 + AAQ._ Users could control their own learning pace by clicking on
the lecture outline. They could look for specific topics by using keyword search or AAQ.

_Treatment 4: A99 + AAQ + Content._ Links to more examples of DD were provided at
the bottom of the lecture outline.

_Treatment 5: A99 + AAQ + Content + Quizzes._ This treatment deployed the complete
functional implementation of A99, including pop-up quizzes. Additionally, more
examples of DD were also provided in this treatment.

The experimental procedure followed the pretest-training-posttest procedure in the
2002 experiment, except that there were only two training sessions in this experiment.
The session on “Heuristics” was removed and its contents were allocated to the
remaining lectures. The two sessions had about the same duration: one and a half hours.
The DD accuracy pre-test and post-test examples, however, were modified carefully,
since the previous results indicated that there could be some problems with the test
questions. The experimental design is summarized in Table 3.

<table>
<thead>
<tr>
<th>Treatment Conditions</th>
<th>Video Only</th>
<th>Linear A99</th>
<th>A99 + AAQ</th>
<th>A99 + AAQ + content</th>
<th>A99 + AAQ + content + quizzes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intro</td>
<td>Pre-tests</td>
<td>Pre-tests</td>
<td>Pre-tests</td>
<td>Pre-tests</td>
<td>Pre-tests</td>
</tr>
<tr>
<td></td>
<td>Instruction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-tests</td>
<td>Post-tests</td>
<td>Post-tests</td>
<td>Post-tests</td>
<td>Post-tests</td>
<td>Post-tests</td>
</tr>
<tr>
<td>Usability Questionnaire</td>
<td>Pre-tests</td>
<td>Pre-tests</td>
<td>Pre-tests</td>
<td>Pre-tests</td>
<td>Pre-tests</td>
</tr>
<tr>
<td></td>
<td>Instruction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cues</td>
<td>Post-tests</td>
<td>Post-tests</td>
<td>Post-tests</td>
<td>Post-tests</td>
<td>Post-tests</td>
</tr>
</tbody>
</table>

Table 3. Experimental Design for Experiment 2
5.2.2.1 Results

For both the introductory and the cues training sessions, as well as across sessions, there was an overall significant improvement on both the DD knowledge test \((t(176)=17.53, p < .001)\) and the DD accuracy test \((t(174)=3.88, p < .001)\) (George et al. 2004a). Therefore, our prediction that training can improve humans’ DD capability was supported. An ANOVA test with planned contrast analysis for knowledge pre-tests and post-tests found significant differences between treatments 4 and 5 for both the Intro \((p = 0.090, \text{one-tail})\) and the Cues \((p = 0.051, \text{one-tail})\) training sessions. It indicated that quizzes did improve the learning of DD knowledge. For DD accuracy tests, the contrast analysis discovered that for the Intro session alone there was a significant difference between treatments 4 and 5 \((p = 0.033, \text{one-tail})\), as well as between treatment 1 and the other four groups \((p = 0.035, \text{one-tail})\). It indicated that the A99 system did improve DD accuracy compared to watching a single video. The results also indicated that the pop-up quiz was indeed an important feature contributing to improve the detection accuracy. However, surprisingly, we did not see any difference among the condition 2, 3, and 4, even though these conditions provided very different functions to users. Did it mean that the new functions (the search tools and more examples) did not contribute at all to the training effectiveness? However, the interactivity and the examples were the functions that the users actually requested in the previous usability study. Again, the information from usability testing provided answers into what happened in the experiment.
5.2.3. Usability Testing

Because the system was modified considerably and many new functions were added, the usability testing was conducted again to evaluate the utility of this new version of the system. With the experiences from the study phase I, we also expected that the results from the usability testing could help us interpret the results from the field experiment. Since the Intro lecture had fewer contents than the Cues lecture, the usability questionnaire was given after the post-test at the end of the first session of the experiment to prevent participants from getting fatigued. The usability questionnaire was modified according to the system changes, although we still focused on similar usability attributes such as ease of use, satisfaction, effectiveness/usefulness, and so on. In this version of the questionnaire, there were twenty nine closed-ended questions rated on a 7-point Likert-type scale. Seven questions were about certain system functions that were only available in certain treatment groups. Therefore an eighth choice “N/A” (not applicable) was added to deal with this situation. Six open-ended questions were given to users to let users describe any problems they experienced in using this system, as well as their preferences about the system.

5.2.3.1 Results

In brief, the usability attributes achieved positive responses across all conditions, with most means close to or more than 5 (“Slightly agree” – “Completely agree”) (see (Cao et al. 2004b) for more details). This indicated that the participants agreed that the A99 system had good usability even with just partial functions. Most of the items related to specific features of A99 (e.g. synchronized multimedia presentation and the self-paced
learning method) received positive responses and were viewed as effective in helping users learn to detect deception. The newly added function, pop-up quizzes, was also rated as very helpful in supporting learning, which strongly supported the findings from the experiment. However, for the three items measuring users’ attitudes towards the search tools, almost all users selected “N/A”, which meant that they either did not know the functions were available to them, or they knew the functions existed but never used them. After carefully analyzing the users’ responses to the open-ended questions, we realized that most users actually had no time to use these new functions at all, because of the time constraint on the lecture and because they were not forced to use these functions (unlike the pop-up quizzes, which were mandatory to be taken during the training). This finding made us realize that even though we designed the condition 2, 3, and 4 to be different, in the real experiment the conditions were actually the same. The design of our research method was somewhat flawed. A follow up study of learning theory literature let us realize that students might not have the motivation to interact with the virtual mentor in the introductory learning phase. Therefore, when designing the experiment to study how virtual interactions may affect the learning effectiveness, another important factor – learning phases – needs to be considered.

In addition, five composite variables that represent the major usability attributes were extracted by a principal components factor analysis from 22 close-ended items that are applicable to all five conditions. They are: Perceived Learning Effectiveness, General Satisfaction, Audio/Video Quality Satisfaction, Ease of use/Learning to use, and Comparison with traditional classroom learning. Results from ANOVA tests with
planned contrasts revealed that the Perceived Learning Effectiveness in all the A99
groups (treatment 2, 3, 4, and 5) was significantly better than in the video only group (t =
1.697, p = .046, one-tailed). Both the Perceived Learning Effectiveness (t = 1.730, p
= .043, one-tailed) and the General Satisfaction (t = 2.366, p = .010, one-tailed) in the
treatment 5 were significantly better than the other A99 groups with partial functions
(treatment 2, 3, and 4). All the other variables were statistically equal across all
conditions (see (Cao et al. 2004b) for details). These findings were consistent with the
results from the experiment, and re-confirmed our conclusion: the mistakes in the
experimental design, especially the exclusion of the learning phase factor and inadequate
time to complete the training, resulted in the unexpected experimental results.
6. INVESTIGATING THE LVM MODEL: THE LVM STUDY

The two A99 studies helped us form the LVM model for interactive e-learning. The findings from these two studies indicated that: 1) interaction is desired in e-learning systems; 2) even provided in e-learning systems, the virtual interaction functions may not be used in the learning process. There must be some other factors such as learning phases and students’ characteristics that affect students’ learning behavior and the overall learning outcomes. Grounded on these findings and learning theories discussed in Chapter 2, we proposed the LVM model that is described in Chapter 3. We then implemented a new prototype e-learning system (named the LVM prototype) based on the LVM model. Again, a multi-methodological research study was conducted to investigate the propositions we made about the LVM model. In this chapter, we describe in detail the new prototype implementation, a series of studies designed to test the LVM model and evaluate the prototype, and the results of those studies.

6.1. System Development

The Agent99 Trainer II used CD-ROM delivery because of the strict military security requirements. However, many disadvantages of this delivery method were exposed during our experimentation. First of all, the capacity of CD-ROM is very limited for multimedia-based lectures: one CD-ROM can only contain one lecture in normal length (around 1 hour). Secondly, the QA function is too complicated to be implemented in local computers so it still requires a remote server. However the communication between the local program and the remote server is not as flexible as using the client/server
architecture. Finally and most importantly, we found that it is not convenient to update the learning materials if using CD-ROM delivery. Therefore, considering that our final goal is to study e-learning for general situations in education and professional training, we returned to the Web-based client/server architecture. Although for some situations like military environments the severe security restrictions may require specially arranged servers, the flexibility of Web-based implementation is more beneficial for most situations.

The implementation of the LVM prototype was similar to the Agent99 Trainer I, except that 1) the examples were integrated into the virtual classroom (instead of having a separate module which may cause confusion in navigation); and 2) more interactive functions were added and integrated into the virtual classroom. The system architecture was still as depicted in Figure 10. All system modules were implemented using HTML, Java Script, VB script, and Java Servlet. As shown in Figure 21 - Figure 24, the user interface design was more integrated and the virtual interaction functions better simulated real classroom interactions.

Figure 21 is the screenshot of the LVM prototype virtual classroom with full interactive functions. The Web-based user interface was divided into four cells or sections: 1) a video display of the instructor, 2) a PPT slide associated with the current video segment, 3) a text note that 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 was synchronized. Furthermore, students might interact with the LVM system in two ways. First, each topic in the outline was directly linked to the relevant video segment,
allowing students to click on any link to review a specific topic. The outline could 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 was provided for students to ask questions. Once a question was submitted, a new window with the four-cell design would pop up to present the potential answers (Figure 22). The answer video segment determined to be the “best fit” would be automatically played, with its associated PPT slide and text note appearing in the other cells. Therefore students could immediately watch and hear the virtual mentor talking back. In the answer window, the lecture outline was replaced with the list of answer topics. If students were not satisfied with the first answer, they could click on the links in this answer list to view the other video answers.

Figure 21. LVM - Virtual Classroom with Full Interactive Functions
Besides the interactions initiated by students, the virtual mentor was also able to initiate the interactions with students automatically. Similar to the Agent99 Trainer II, pop-up quizzes were implemented in this LVM prototype (Figure 23). However, a QA box was added into the quiz feedback page (Figure 24) so that if students found their answer not correct, they could ask the virtual mentor a question immediately.
Figure 23. LVM – Pop-up Questions

Figure 24. LVM – Feedback to Pop-up Questions
In addition, to study the relationship between interactive functions and learning outcomes, we built different versions of the LVM prototype with different types or amounts of interactive functions. They are depicted in Figure 25, Figure 26, and Figure 27. The version shown in Figure 27 had the least amount of interactive functions. It was mainly a linear playback of the entire lecture with synchronized cells. The only way that students might navigate in the lecture was to use the fast forward and backward buttons in the Real player console.

Two video lectures were created to be used in this prototype. One (about 27 minutes) is a lecture about relational normalization, a concept in database management. This lecture was used in a pilot study. The other (about 40 minutes) is a lecture about computer security, and was used in the final research study.

Figure 25. LVM - Virtual Classroom without Outline and Navigation Buttons
Figure 26. LVM - Virtual Classroom without QA

Figure 27. LVM - Virtual Classroom with Linear Playback
6.2. Experimentation and Observation

When investigating the LVM model (Figure 9), we conducted a field experiment in conjunction with observations such as survey and interviews. To explore the relationship among the learning outcomes and the controllable factors such as system interactivities and learning phases, an experiment was conducted as a longitudinal pretest-posttest comparison between control and treatment groups (Table 4). Since we were mostly interested in the relationship between the system interactivities and the learning outcomes, we arranged five treatment groups where participants used the LVM prototype with varying system interactivities, and one control group where a real mentor taught the participants face-to-face. As shown in Table 4, system interactivities were increased from treatment group 1 to treatment group 5 (left to right as in Table 4; IN means students could only watch the instruction sequentially, Outline means that students could click on the links in the outline to change topic, QA means students could ask questions, Quiz means students received pop-up questions in the middle of their learning process; see Figure 21-Figure 27 for screenshots of the user interface for different treatment groups). As mentioned in Chapter 3, we also think that learning phases may affect students’ interactive learning activities and learning outcomes. Therefore, we designed two sessions in this experiment. The first session, class session, is the introductory learning phase where students were introduced to the lecture content the first time. The second session, assignment session, is the advanced learning phase where students were required to review the lecture and complete an assignment about the lecture content. There were two weeks between the two sessions and participants were required to attend both
sessions. Therefore we can compare the same student’s behaviors and learning outcomes in the two different learning phases. Pre-test and post-test were deployed to each session.

<table>
<thead>
<tr>
<th>Treatments</th>
<th>Real Mentor</th>
<th>VM – IN only</th>
<th>VM – IN + Outline</th>
<th>VM – IN + QA</th>
<th>VM – IN + Outline + QA + Quiz</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Class session</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-tests (45 min)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning (50 min; including 10 min review time)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-tests (30 min)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Assignment session</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delayed post-test (15 min)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finishing an assignment (50 min)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-tests (30 min)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Experimental Design

To explore the relationship among the learning outcomes, learning activities, and students’ characteristics such as learning motivation/strategy, a correlational study was conducted in conjunction with the experiment using surveys deployed before and after the experiment’s pre/post-test. Students’ characteristics were measured using a pre-experiment survey; and their perceptions about the system and the learning effectiveness were collected using a post-experiment survey.

Besides the quantitative data collected in the experiment and the surveys, we also collected a large amount of qualitative data using open-ended questions in the post-experiment survey. In addition, some students volunteered to complete a semi-structured interview after the experiment. However, only survey responses were analyzed in this dissertation.
6.2.1. Study Design

6.2.1.1 Participants and Treatment Groups

Participants (N = 158) were undergraduate students (59% male, 41% male) recruited from an Information Systems (IS) course at a southwestern university. They volunteered to participate in this research study as an alternative to 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 fulfill credits for this assignment as long as they actively participated in all experimentation sessions. Sixty nine percent of the students reported their GPA as between 3.0-4.0, while thirty one percent of the students reported their GPA as between 2.0-3.0.

6.2.1.2 Procedures

Participants signed up for this research study through a website and they could choose one from six available time slots for attending. The six time slots were randomly assigned to the six groups introduced above, so students were randomly pre-assigned to one of the treatment groups or the control group (without their own awareness of the condition of the group) when they entered the computer lab for the experiment. All groups used the same computer lab for the research study, including the Real Mentor group.

The study began with the author consenting students about the study design and procedures. Students in each group were required to finish both two sessions of learning. They were not allowed to switch to a different group in the second session after two weeks. As stated earlier, the first session was designed to be introductory, non-task-
oriented learning. In this session, participants first filled out a pre-experiment survey (30min total; including demographic information and surveys that measuring students’ learning motivation/strategy/styles) and then completed a pre-test that had 10 multiple-choice questions about computer security (15 min). Then they received the computer security lecture in their particular treatment conditions (50 min; with human instructor or using the LVM prototype). Finally they completed a post-test, which had the same question format as in the pre-test but the order of both the questions and the answer choices were alternated (15 min). After this post-test they answered a post-experiment questionnaire (15 min; including questions about system usability, perceived learning effectiveness, and other factors). The participants’ activities in using the LVM system were recorded into 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 this second session and they were required to learn with specific tasks – finishing an assignment. They first took a 15-minute delayed posttest on their knowledge about computer security, then completed an open-ended assignment in 50 minutes with the help of either asking instructors directly (in the control group) or using the system (in the treatment groups), and finally completed another posttest (15 min). At the end of this session they finished the same post-experiment questionnaire as in the first session (15 min). Again, during the 50 minutes assignment session, the participants’ activities in using the training system were
recorded into a system log file, and the activities of the participants in the “Real Mentor”
group were videotaped for future analysis.

After the first session, an average of three or four individuals in each group
volunteered to participate in a semi-structured face-to-face interview with the author in a
different location. Each interview lasted from 15 minutes to 30 minutes. All interviews
were recorded for future analysis. After the second session, because of the time
constraints, email interviews were used to replace the face-to-face interviews. Themes of
the interview questions were reinforcements of the open-ended questions asked in the
post-learning questionnaire. Examples of the themes are:

1. Student’s perception of the new technology – like or dislike, why?
2. Student’s detailed learning process;
3. Comparison to their regular learning behaviors;
4. Student’s interpretation of interactivity in learning
5. Student’s expectation for interactive e-learning
6. Student’s understanding of factors for effective learning

Although some responses from the interviews have been used for helping us better
interpret the students’ learning process and the relationship among the different factors
that influence the learning outcomes, in this dissertation we only focus on the qualitative
data from the open-ended questions in the post-learning survey.
6.2.1.3 Measures and Instruments

Independent Variables for the Experiment

1) System Interactivities

Participants were randomly assigned to one of the six groups in which they would have access to different levels of interactivities. We assumed that a real mentor could provide the most interactivities in the learning process and so arranged the Real-Mentor group as our control group. For the other groups, we controlled the system functionalities so that different versions of the system could support different levels of interactivities. The five treatment groups are listed below with system interactivities increasing from treatment group 1 to treatment group 5:

**Treatment 1.** VM - IN only (Figure 27): students could only watch the instruction sequentially;

**Treatment 2.** VM – IN + Outline (Figure 26): students could click on the links in the outline to change topics when watching the instruction;

**Treatment 3.** VM – IN + QA (Figure 25): students could ask questions when watching the instruction;

**Treatment 4.** VM – IN + Outline + QA (Figure 21): students could not only click on the links in the outline to change topics but also ask questions when watching the instruction;

**Treatment 5.** VM – IN + Outline + QA + Quiz (Figure 21 & Figure 23): students could not only click on the links in the outline to change topics but also ask questions.
when watching the instruction; In addition, the system could pop-up questions in the middle of the instruction.

2) Learning Phases

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

Independent Variables for the Correlational Study

1) Learning Motivation/Strategies/Styles

We measured the construct “Learning motivation/strategies/styles” in the LVM using three existing self-report instruments consisting of total 11 scales.

The Motivated Strategies for Learning Questionnaire (MSLQ) was a self-report instrument designed to assess college students’ motivational orientations and their use of different learning strategies for a college course (Pintrich et al. 1993). There were fifteen different scales consisting of a total of 81 items for assessing students’ motivation and learning strategies for general college-level courses. Students rated themselves on a seven-point Likert scale ranging from “not at all true of me” (scale = 1) to “very true of me” (scale = 7). Scale means were used to measure an individual’s score. The scales were designed to be modular so that they could be used together or singly. Therefore, we only selected 6 scales that were relevant to interaction in learning. Each scale measured one independent variable in the “Learning motivation/strategies/styles” construct. Therefore, the six variables were:
• INTRINSIC GOAL ORIENTATION (IG) (four items). This motivation scale measured the degree to which the student perceived herself to be participating in a task for reasons such as challenge, curiosity, and mastery. We thought that students having higher intrinsic goal orientation might have more interactive activities and better learning outcomes in the learner-centered e-learning environment.

• EXTRINSIC GOAL ORIENTATION (EG) (four items). This motivation scale measured the degree to which the student perceived herself to be participating in a task for reasons such as grades, rewards, performance, evaluation by others, and competition. Because of the highly learner-centered nature of the e-learning environment, we proposed that students high in extrinsic goal orientation might engage less in the learning process and have less interactive activities with the virtual mentor, and thus have worse learning outcomes.

• TASK VALUE (TV) (six items). This motivation scale referred to the student’s evaluation of how interesting, how important, and how useful the task was. High task value should lead to more involvement in one’s learning and better learning outcomes.

• ORGANIZATION (ORG) (four items). This scale was about the student’s strategies on selecting appropriate information and constructing connections among the information to be learned. Organizing is an active, effortful endeavor, and should result in more involvement in one’s learning and better learning outcomes.

• METACOGNITIVE SELF-REGULATION (MSR) (twelve items). This scale measured the student’s ability to control cognition and engage in self-regulation. For example, students reported how well they planned, monitored and regulated learning,
such as self-testing and questioning. High score in this scale meant that the student was more self-regulated and thus should result in better performance in the e-learning environment.

- HELP SEEKING (HS) (four items). This scale concerned how students manage the support of others in learning. Good students know when they don’t know something and are able to identify someone to provide them with some assistance. High score in this scale indicated that the student was willing to seek help in learning and thus should result in more interactions with the virtual mentor and finally had better learning outcomes.

Besides the students’ learning motivation and strategies, we were also interested in their learning styles. The learning style model we selected was described in (Felder 1993). This model identified students’ learning styles on four dimensions (active/reflective, sensing/intuitive, visual/verbal, and sequential/global). The students’ preferences on these four dimensions were measured by an Index of Learning Styles (ILS) instrument. The ILS results provided an indication of possible strengths and possible tendencies or habits that might lead to difficulty in academic settings. However, in this dissertation study, we only used three dimensions that we believed to be more relevant to the issues of interaction in learning. Each dimension contained eleven items and the each item had two response choices that the student needed to choose from. ILS authors provided a scoring sheet to calculate the score for each dimension. Again, items in each dimension measured one independent variable in the “Learning motivation/strategies/styles” construct and three variables are:
• **ACTIVE AND REFLECTIVE LEARNERS (ACT):** “Active learners tend to retain and understand information best by doing something active with it—discussing or applying it or explaining it to others. Reflective learners prefer to think about it quietly first. Active learners tend to like group work more than reflective learners, who prefer working alone.” (Felder 1993) This description of active and reflective learners indicated that sitting through lectures without physically doing anything but taking notes was particularly hard for active learners. So if the student was an active learner (higher score on this scale), she would need more interaction in the learning process to be successful in an e-learning environment.

• **SENSING AND INTUITIVE LEARNERS (SNS):** “Sensing learners tend to like learning facts, intuitive learners often prefer discovering possibilities and relationships. Sensors tend to be patient with details and good at memorizing facts; intuitors don’t like “plug-and-chug” courses that involve a lot of memorization and routine calculations.” (Felder 1993) This description indicated that if the student was an intuitive learner (lower score on this scale), she would need more interaction in the learning process to be successful in an e-learning environment.

• **SEQUENTIAL AND GLOBAL LEARNERS (SEQ):** “Sequential learners tend to gain understanding in linear steps, with each step following logically from the previous one. Global learners tend to learn in large jumps, absorbing material almost randomly without seeing connections, and then suddenly ‘getting it’.” (Felder 1993) This description indicated that if the student was a global learner (lower score on this
scale), she would need more interaction in the learning process to be successful in an
e-learning environment.

The MSLQ and ILS provided numerous scales measuring students’ individual
learning strategies and styles. However, these scales still might not explain the students’
individual differences in different learning phases. The Approaches to Learning and
Studying (ALS) Inventory (Entwistle 1998) offered a possible solution to this problem.
This instrument measured a student’s strategic approach to studying and had two sub-
scales that we thought to be able to assess whether students’ general learning strategies
would be fitful for different learning phases. These two sub-scales measured the final two
independent variables in the “Learning motivation,strategies/styles” construct, as shown
below. Each sub-scale consisted of 8 items. Students rated themselves on a five-point
Likert scale ranging from “disagree” (scale = 1) to “agree” (scale = 5). An individual’s
score on a scale was conducted by taking the mean of the items that made up that scale.

- **DEEP APPROACH (DEEP):** A high score on this scale meant that the student tended
to develop understanding of the topic under study, and tended to show active
engagement and interest in their studies. Students scoring higher on this scale were
predicted to have better learning outcomes than those who had lower score on this
scale in the second session in our study, which required more advanced, deep learning.

- **SURFACE APPROACH (SURFACE):** A high score on this scale meant that the
student tended to memorize information to meet external (assessment) requirements
without critical thinking and understanding. We predicted that students scoring higher
on this scale would have better learning outcomes than those who had lower score on
this scale in the first session in our study, because the first session required more
introductory, surface learning.

2) Learner Capabilities

The learner capability was measured by participants’ self-reported GPA, with three
ranges: 1.0-2.0, 2.0-3.0, 3.0-4.0.

Dependent Variables

1) Learning Activities

This variable was only 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: asking a question, switching topic using the
outline, switching topic using the navigation buttons. This variable was measured by the
total number of the student’s interactive activities (#INT) during a session. The advanced
learning phase and more system interactivities were expected to trigger more interactive
learning activities. However, this dependent variable might also correlate with the other
two dependent variables about learning outcomes in that the more interactive activities,
the better learning outcomes. We conducted a correlation analysis to explore the
relationship among them.

2) Learning Performance

The student’s objective learning performance was measured by her percentage
accuracy score on the post-tests (10 multiple-choice questions on the student’s
knowledge of computer security), as compared to the score of pre-test. 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.

3) 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). 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 items that made up that scale. A high score on this scale meant that the student thought she learned effectively in this study.

4) Perceived System Usability

The perceived system usability was measured by a scale named System Usability Scale (SUS) (Brooke 1996). This scale included 10 items and was included in the post-experiment questionnaire. Again, an individual’s score on the scale was computed by taking the mean of the items that comprised that scale. A high score on this scale meant that the student thought the system was easy to learn to use and easy to use. This variable was only applicable to the five treatment groups.

5) Self-report Interactivity

This measure (INTER) was developed by the author. It assessed whether students were satisfied with their interaction with the virtual mentor (or real mentor in the control group) or not. The four-item scale on the post-experiment questionnaire was averaged to produce a mean score for each individual. A high score on this scale meant that the
student felt that they had interactions with the (virtual) mentor and enjoyed the (virtual) interactions.

6) Learning System Preference

This variable (LSP) assessed the student’s preference on learning methods. It was measured by a self-developed four-item scale in the post-experiment questionnaire. 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 preferred the e-learning method other than the traditional instructor-led, classroom learning.

Finally, there were twelve more items in the post-experiment questionnaires that measured the student’s perception on specific system functionalities. These items were treatment-group-specific (some functions were not available in some treatment groups) and so were only used for descriptive purposes.

6.2.1.4 Hypotheses for the Experiment

Based on the preceding literature review and the LVM model described in Chapter 3, we proposed the following hypotheses about the causal relationship among the system interactivity, leaning phases, learning activities, and the learning outcomes.

**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 the overall learning process.

**Hypothesis 2:** Higher levels of system interactivity result in better learning outcomes in the advanced learning phase, as well as the overall learning process.
Hypothesis 3: Students will engage in more interactive learning activities in the advanced learning phase than in the introductory learning phase.

Hypothesis 4: Increasing the level of system interactivity results in a greater improvement in learning outcomes in the advanced learning phase than in the introductory learning phase.

For these hypotheses, learning outcomes are measured through the following dependent variables: learning performance, perceived learning effectiveness (EFFECT), perceived system usability (SUS), self-report interactivity (INTER), and learning system preference (LSP).

6.2.1.5 Hypotheses for the Correlational Study

To explain the relationship between the learner characteristics and the learning outcomes, we proposed the following research hypotheses.

Hypothesis 5: Students who are more a) Intrinsic Goal Oriented (high score on IG); b) with high Task Value (high score on TV); c) Organized (high score on ORG); d) Metacognitive Self-Regulated (high score on MSR); e) Help Seeking (high score on HS); f) Active Learners (high score on ACT) engage in more learning activities, perform better, have higher perceived learning effectiveness, report more interactivity and show greater preference for the system in all learning phases.

Hypothesis 6: Students who are less a) Extrinsic Goal Oriented (low score on EG); or who are b) Intuitive Learners (low score on SNS); c) Global Learners (low score on SEQ) engage in more learning activities, perform better, have higher perceived learning
effectiveness, report more interactivity and show greater preference for the system in all
learning phases.

**Hypothesis 7**: Students who take Deep Approach in learning (high score on DEEP) engage in more learning activities, perform better, have higher perceived learning effectiveness, report more interactivity and show greater preference for the system in the *advanced* learning phase.

**Hypothesis 8**: Students who take Surface Approach in learning (high score on SURFACE) engage in more learning activities, perform better, have higher perceived learning effectiveness, report more interactivity and show greater preference for the system in the *introductory* learning phase.

**Hypothesis 9**: Students who engage in more learning activities will perform better, have higher perceived learning effectiveness, report more interactivity and show greater preference for the system.

**Hypothesis 10**: Students who perform better also have higher perceived learning effectiveness.

### 6.2.2. Analysis and Results

Both quantitative and qualitative data were collected in this study. Quantitative data were examined by reliability test, factor analysis, repeated measures analysis and/or ANCOVA, and multiple regressions. All statistical analysis was conducted by using SPSS 10.0. Qualitative data obtained from the open-ended questions in the questionnaires were analyzed using theme-based content analysis approach. This chapter reports in detail the results of these analyses and the implications of the results.
6.2.2.1 Reliability Test and Factor Analysis

A reliability test was conducted to see how reliable the questionnaire instruments were in our study. The same post-experiment questionnaire was given to students after each learning session. Therefore the reliability test was conducted for each session to see whether there were substantial differences.

Table 5 lists the reliability for various constructs in the post-experiment questionnaire for each session. Reliabilities for all constructs were above .81, which was an indication of good reliability.

Table 5. Reliabilities for the Constructs in the Post-Experiment Questionnaire

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cronbach Alpha</th>
<th>Cronbach Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Session 1 (N = 158)</td>
<td>Session 2 (N = 152)</td>
</tr>
<tr>
<td>EFFECT</td>
<td>.88</td>
<td>.91</td>
</tr>
<tr>
<td>SUS</td>
<td>.98</td>
<td>.97</td>
</tr>
<tr>
<td>INTER</td>
<td>.82</td>
<td>.88</td>
</tr>
<tr>
<td>LSP</td>
<td>.92</td>
<td>.95</td>
</tr>
</tbody>
</table>

Table 6 lists the reliability for various variables in the pre-experiment survey. Reliabilities for most variables were above .70, which was acceptable.

Table 6. Reliabilities for the Constructs in the Pre-Experiment Survey

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cronbach Alpha (N = 158)</th>
<th>Variable</th>
<th>Cronbach Alpha (N = 158)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IG</td>
<td>.75</td>
<td>ACT</td>
<td>.61</td>
</tr>
<tr>
<td>EG</td>
<td>.70</td>
<td>SNS</td>
<td>.73</td>
</tr>
<tr>
<td>TV</td>
<td>.88</td>
<td>SEQ</td>
<td>.38</td>
</tr>
<tr>
<td>ORG</td>
<td>.66</td>
<td>DEEP</td>
<td>.73</td>
</tr>
<tr>
<td>MSR</td>
<td>.72</td>
<td>SURFACE</td>
<td>.82</td>
</tr>
<tr>
<td>HS</td>
<td>.23</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
However, the variable HS from the MSLQ and the variable SEQ from the ILS had very low reliability scores. The lower scores for the variables from the ILS were still understandable because variables in this survey were not developed by directly summing up the scales. A special scoring procedure was applied for calculating the score of each variable. Therefore, Cronbach alpha might not be very suitable for testing the reliability of these variables. Even so, the reader is advised to exercise caution in interpreting results about these variables. For the variables from the MSLQ, all variables about learning strategies had relatively low reliabilities. In addition, there were too many variables from MSLQ and this might bring difficulties for correlation analysis and results interpretation. Therefore, we conducted an exploratory factor analysis (EFA) to see whether the number of factors could be deducted and the reliability of the new factors could be acceptable.

*Alpha factor analysis* with varimax rotation was used to extract factors with maximum reliability (Thompson 2004). Using the rule of Eigenvalue greater than 1.0, three factors were extracted (see Figure 28 for the scree plot). The items in the original EG scale was still loaded nicely as a single factor. The items in the original IG and TV scales were loaded together into a new factor (IntrinsicG) and together they described whether students want to learn driven by their own interest in learning or in the subject matter. The items in the original MSR, ORG, and HS scales were loaded together into another new factor (Strategy), describing whether students tend to take a strategic approach in learning.
We tested the reliability of the two new factors and both of them had high Cronbach Alpha scores this time. The updated constructs and their reliability scores were listed in Table 7.

Table 7. Reliabilities for the Updated Constructs in the Pre-Experiment Survey

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cronbach Alpha (N = 158)</th>
<th>Variable</th>
<th>Cronbach Alpha (N = 158)</th>
<th>Variable</th>
<th>Cronbach Alpha (N = 158)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IntrinsicG</td>
<td>.88</td>
<td>ACT</td>
<td>.61</td>
<td>DEEP</td>
<td>.73</td>
</tr>
<tr>
<td>ExtrinsicG</td>
<td>.70</td>
<td>SNS</td>
<td>.73</td>
<td>SURFACE</td>
<td>.82</td>
</tr>
<tr>
<td>Strategy</td>
<td>.81</td>
<td>SEQ</td>
<td>.38</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For the constructs in the post-experiment questionnaires, we also conducted an EFA to examine their factor loading. Results from both sessions showed that most of the items loaded nicely into the four factors (see Figure 29 and Figure 30 for scree plots) as we expected. Therefore, all original items were included when the composite scores (mean of the individual items) were calculated for the four constructs.
6.2.2.2 Hypotheses Testing

Hypotheses 1 to 4 were tested with repeated measures Analyses of Variance (ANOVA). Multivariate analyses were initially conducted on sets of related variables, followed by simple effect tests on the individual dependent variables. When analyzing

Figure 29. Scree Plot for the Post-Experiment Items Analysis (Session 1)

Figure 30. Scree Plot for the Post-Experiment Items Analysis (Session 2)
the data about learning performance, we also conducted Analysis of Covariance (ANCOVA) on posttest data with the pretest scores used as the covariate. This approach is recommended by (Bonate 2000) because of the greater statistical power and lack of bias that is achieved with ANCOVA when analyzing pretest-posttest data.

Results about Interactive Learning Activities

Table 8 lists the means and standard deviations of students’ interactive learning activities in each session. As mentioned earlier, this variable was only measured for the five treatment groups where all students’ learning activities were recorded into a system log file. However, in the first group, the “IN only” group, the students had no access to any interactive system functions such as QA and outline. Therefore, only four groups had interactive learning activities recorded in system logs; analysis was thus based on these four groups.

Table 8. Means and Standard Deviations of Interactive Learning Activities

<table>
<thead>
<tr>
<th>Group</th>
<th>Session 1 Activities N</th>
<th>mean (std)</th>
<th>Session 2 Activities N</th>
<th>mean (std)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 (IN + Outline)</td>
<td>25</td>
<td>6.44 (6.01)</td>
<td>25</td>
<td>19.28 (10.50)</td>
</tr>
<tr>
<td>3 (IN + QA)</td>
<td>24</td>
<td>2.92 (2.90)</td>
<td>23</td>
<td>27.54 (8.71)</td>
</tr>
<tr>
<td>4 (IN + QA + Outline)</td>
<td>28</td>
<td>9.04 (4.95)</td>
<td>28</td>
<td>25.04 (14.14)</td>
</tr>
<tr>
<td>5 (IN + QA + Outline + quizzes)</td>
<td>30</td>
<td>6.50 (7.58)</td>
<td>30</td>
<td>22.20 (13.41)</td>
</tr>
<tr>
<td>Total</td>
<td>158</td>
<td>6.35 (6.05)</td>
<td>152</td>
<td>23.46 (12.27)</td>
</tr>
</tbody>
</table>

Repeated measures analysis indicated that students’ interactive learning activities were different for the factor Session, \(F(1, 103) = 204.776, p < .005, \text{ partial } \eta^2 = .665 \); and interaction Session*Treatment, \(F(3, 103) = 4.097, p = .009, \text{ partial } \eta^2 = .107 \) (see Figure 31). The interaction indicated that although all groups had more interactive learning
activities in the advanced learning session, the increase was different among groups. A planned, reverse Helmert contrast analysis was conducted on the Treatment factor. This contrast analysis compares the mean of the first or last group to the mean of the other groups, and then conducts similar comparisons among the remaining groups. In this case it compared the mean of group 5 (IN+QA+Outline+Quiz) with the mean of group 2 (IN+QA+Outline), 3 (IN+QA), and 4 (IN+QA+Outline); then mean of group 4 with the mean of group 2 and 3; and finally mean of group 3 with mean of group 2. Results indicated that students in group 3 had significantly more interactive activities than group 2 in session 2 (p = .018); however the more system interactivities in group 4 and 5 did not significantly increase further the actual learning activities in this advanced session. For overall interactive activities across both sessions, group 4 had significantly more interactive activities than the mean of group 2 and 3 (p = .079); however group 5 did not have significantly more interactive learning activities than the other groups.

Figure 31. Means of Interactive Learning Activities in Two Sessions
Therefore, hypotheses 1 and 3 received mixed results. A follow up pair-wise comparison on the interactive activities in the session 2 also did not show significant differences among group 3, 4 and 5. After carefully examining the study procedure, we found that the students’ activities of answering the pop-up questions in the first session were not recorded in the system logs and the pop-up questions were designed not to show up in the second session. Therefore, it was reasonable that group 4 and 5 had similar activity patterns. Since the major difference between group 2 and group 3, 4, 5 was then the QA function, we concluded that it was the QA function that triggered more interactive learning activities in the advanced learning session.

Another interesting observation was that, although we assumed that the real mentor would be able to provide the most interactivities, in both sessions there were very few interactions between the students and the instructor in the classroom. In the first session, there were only two students asking three questions at the end of the lecture; in the second session, there were only four students asking the instructor questions when doing the assignments. This was obviously different from what we expected. Our qualitative results from the semi-structured interviews after the experiment offered some explanation. 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 students’ psychological barrier and enable more direct question and answering processes in learning.

Results about Learning Performance

Table 9 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.

<table>
<thead>
<tr>
<th></th>
<th>Session 1 (Class Session)</th>
<th>Session 2 (Assignment Session)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
<td>Posttest</td>
</tr>
<tr>
<td></td>
<td>Mean (std)</td>
<td>mean (std)</td>
</tr>
<tr>
<td>IN only</td>
<td>26</td>
<td>28.8 (16.1)</td>
</tr>
<tr>
<td>IN + Outline</td>
<td>25</td>
<td>29.2 (17.1)</td>
</tr>
<tr>
<td>IN + QA</td>
<td>24</td>
<td>26.3 (15.0)</td>
</tr>
<tr>
<td>IN + QA + Outline</td>
<td>28</td>
<td>32.5 (21.0)</td>
</tr>
<tr>
<td>IN + QA + Outline + quizzes</td>
<td>30</td>
<td>29.0 (14.7)</td>
</tr>
<tr>
<td>Real Mentor</td>
<td>25</td>
<td>30.4 (21.7)</td>
</tr>
<tr>
<td>Total</td>
<td>158</td>
<td>29.4 (17.6)</td>
</tr>
</tbody>
</table>

A $6 \times 2 \times 2$ repeated measures analysis was conducted to test the hypotheses. Since the F-statistic for the interaction between the treatment factor and the pretest-posttest factor is identical to the F-statistic for the treatment main effect with a one-way ANOVA on gain scores (Bonate 2000), we used the interaction F ratio, not the main effect F ratio, for testing the treatment main effect.

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 $\eta^2 = .124$ (see Figure
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. This is quite understandable because in session 2 there was less room for improvement. However, no significant treatment effect was found through the repeated measures analysis. All groups had similar patterns in test score changes.

As suggested in (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 the overall
learning, we conducted two ANCOVA tests respectively, with the session 1 pretest as the 
covariate and the posttest scores as the dependent variable for each session.

The results were indeed different from the results from the repeated measures analysis. 
There were significant differences among groups in both session 2, F(5, 145) = 2.274, p 
= .05, partial η² = .073, see Figure 33, and overall, F(5, 145) = 2.409, p = .039, partial η² 
= .077 (see Figure 34). In session 2, reverse Helmert contrasts 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. Based on the findings 
from both the repeated measures analysis and the ANCOVA test, we concluded that there 
was no significant difference among system groups found in either the second session or 
the overall learning process. However, the real mentor group, which was assumed to have 
the highest learning performance gain in both session 2 and overall, 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 the hypothesis 2 might still 
be partly supported. However, more studies are needed for retesting this hypothesis.
At SESSION = 2

TREAT ID IN ORDER
- Group 1 IN only
- Group 2 IN+Outline
- Group 3 IN+QA
- Group 4 IN+QA+
- Group 5 IN+Outline+QA+Quiz
- Group 6 Real Mentor

Figure 33. Means of Test Scores (%) in Session 2

PREPOST

TREAT ID IN ORDER
- Group 1 IN only
- Group 2 IN+Outline
- Group 3 IN+QA
- Group 4 IN+QA+
- Group 5 IN+Outline+QA+Quiz
- Group 6 Real Mentor

Figure 34. Means of Test Scores (%) from the Beginning to the End
Results about Perception/Satisfaction

Table 10 lists the means and standard deviations of the four variables from the self-report post-experiment survey in each session. Among these four variables, SUS and LSP were only applicable to the five treatment groups where students were learning with the virtual mentor. Therefore we conducted two multivariate repeated measures analyses. One had EFFECT and INTER as dependent variables and the independent variable Treatment had six groups. Another had all four variables as dependent variables and the independent variable Treatment only included the five system groups.

Table 10. Means and Standard Deviations of Self-Reported Learning Outcomes

<table>
<thead>
<tr>
<th></th>
<th>Session 1 (Class Session)</th>
<th>Session 2 (Assignment Session)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EFFECT mean (std)</td>
<td>INTER mean (std)</td>
</tr>
<tr>
<td>IN only</td>
<td>26 3.64 (.50)</td>
<td>2.68 (.82)</td>
</tr>
<tr>
<td>IN + Outline</td>
<td>25 3.56 (.53)</td>
<td>2.91 (.77)</td>
</tr>
<tr>
<td>IN + QA</td>
<td>24 3.68 (.68)</td>
<td>3.04 (.76)</td>
</tr>
<tr>
<td>IN + QA + Outline</td>
<td>28 3.47 (.68)</td>
<td>2.94 (.84)</td>
</tr>
<tr>
<td>IN + QA + Outline + quizzes</td>
<td>30 3.65 (.58)</td>
<td>3.06 (.66)</td>
</tr>
<tr>
<td>Real Mentor</td>
<td>25 3.86 (.72)</td>
<td>3.98 (.55)</td>
</tr>
<tr>
<td>Total</td>
<td>158 3.64 (.62)</td>
<td>3.10 (.83)</td>
</tr>
</tbody>
</table>

Hypotheses about EFFECT (part of hypotheses 2 and 4) received little support from both analyses. The multivariate analysis on the EFFECT and INTER measures with six groups showed that students’ perceived learning effectiveness and perceived interaction with the (virtual) mentor were significantly different between the two sessions, Wilk’s λ.
= .961, F(2, 143) = 2.938, p = .056, η2 = .039. The follow up univariate analyses revealed that students’ perceived learning effectiveness in session 2 was significantly less than in session 1, F(1,144) = 5.914, p = 0.016 (Figure 35). The analysis with five groups yielded similar results (univariate analysis on EFFECT, F(1,121) = 3.399, p = 0.068). We did not find significant differences among treatment groups in session 2. In addition, even though group 4 (IN+QA+Outline) seemed to have increased EFFECT across sessions, a comparison on the difference scores failed to find significant difference among the groups on how they changed from session 1 to session 2. 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 contents 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, there were no significant differences among the five treatment groups.
Hypotheses about INTER (part of hypotheses 2 and 4) also received little support. In both analyses (with six groups or with five groups), the difference between sessions was not significant across groups (see Figure 36). However, analysis with six groups showed significant difference among treatment groups (between-subject main effect, F(5,144) = 12.490, p < .005). In addition, both analyses showed a significant interaction for Session*Treatment, F(5,144) = 1.977, p = .082 in six group analysis; F(4,121) = 2.160, p = .078), which meant the change of perceived interactivity from session 1 to session 2 was different among groups. 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 with six-group analysis; p = .002 with five-group analysis) and in the overall learning process (p = .010 with six-group analysis; p = .013 with five-group analysis). However, group 4 and 5 which were supposed to have more system interactivities had either similar or even less perceived interactivities in comparison to group 3. This was contrary to our expectations but quite similar to what happened with the actual interactive learning activities. A Pearson correlation analysis was conducted but no significant correlations were found between the perceived interactivities and the actual active learning activities in both sessions. Therefore, there was no evidence that the students’ perceived interactivities in the advanced learning session were influenced by their actual interactive learning activities. More studies are needed for determining whether it was because QA was the key for different interactive learning activities and thus for different levels of perceived interactivities.

The Session*Treatment interactions were rather complicated for this INTER variable. A Helmert contrast analysis on change scores revealed that the group 3’s change was significantly more than the mean of the group 1’s and group 2’s changes (p = .036). Examining the Figure 36, we found that group 1’s perceived interactivities remained almost the same, while group 2’s perceived interactivities decreased and group 3’s perceived interactivities increased. It is surprising that in the advanced learning session, a system with more interactive functions (outline) actually resulted in decreased perceived learning activities. We offer a possible explanation of this unexpected result based on qualitative data in the next section.
For the hypotheses about SUS (part of hypotheses 2 and 4), the multivariate analysis with five groups on all four measures showed significant differences between the two sessions (Wilk’s $\lambda = .925$, $F(4, 118) = 2.409$, $p = .053$). The follow up univariate analyses revealed that students’ perceived system usability in session 2 was significantly less than in session 1, $F(1,121) = 9.786$, $p = 0.002$ (Figure 37). There was also a significant interaction Session*Treatment ($F(4,121) = 2.237$, $p = .069$), which meant the change of perceived system usability from session 1 to session 2 was different among groups. A reverse Helmert contrast analysis found that the perceived system usability was significantly higher in group 4 than the mean of group 1, 2, and 3 in session 2 ($p = .057$). Also the change of perceived system usability from session 1 to session 2 in group 4 was significantly more than the mean of group 1, 2, and 3 ($p = .006$). Figure 37 shows that
except group 4 which had increased perceived system usability, all the other groups had decreased perceived system usability.

![Means of SUS in Both Sessions](image)

**Figure 37. Means of SUS in Both Sessions**

Finally, for the hypotheses about LSP (part of hypotheses 2 and 4), the univariate analyses revealed that students’ learning system preference in session 2 was not significantly different from in session 1, (Figure 38). Although group 3 and 4 seemed to have increased LSP across sessions, no significant interaction Session*Treatment was found, either.
Results of the Correlational Studies

Table 11 lists the means and standard deviations of the eight learning motivation/strategy/style measures that we collected from the pre-experiment survey. To test hypotheses 5 to 10, we conducted two multiple regression analyses and three correlation analysis.

The first regression analysis had the interactive learning activities be regressed on the eight learning motivation/strategy/styles measures. Results showed that the interactive learning activities in session 1 were not significantly regressed on any of the eight measures. However, the interactive learning activities in session 2 were significantly regressed on one of the learning style measures: ACT, $\beta = .229$, $t(107) = 2.289$, $p = .024$ (see Table 12). Therefore hypothesis 5 was partially supported in the advanced learning
session. More active types of learners (higher score on ACT) would have more interaction in the advanced, task-oriented learning process.

Table 11. Means and Standard Deviation of Learning Motivation/Strategy/Style Measures

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>IntrinsicG</th>
<th>ExtrinsicG</th>
<th>Strategy</th>
<th>ACT</th>
<th>SNS</th>
<th>SEQ</th>
<th>DEEP</th>
<th>Surface</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>mean (std)</td>
<td>mean (std)</td>
<td>mean (std)</td>
<td>mean (std)</td>
<td>mean (std)</td>
<td>mean (std)</td>
<td>mean (std)</td>
<td>mean (std)</td>
</tr>
<tr>
<td>IN only</td>
<td>26</td>
<td>4.43 (1.04)</td>
<td>5.13 (1.00)</td>
<td>3.84 (.89)</td>
<td>2.15 (4.67)</td>
<td>1.77 (5.83)</td>
<td>.65 (2.87)</td>
<td>3.58 (.62)</td>
<td>2.86 (.90)</td>
</tr>
<tr>
<td>IN + Outline</td>
<td>25</td>
<td>4.65 (.98)</td>
<td>5.37 (1.03)</td>
<td>3.95 (.88)</td>
<td>1.40 (4.90)</td>
<td>2.20 (5.60)</td>
<td>.76 (5.13)</td>
<td>3.40 (.55)</td>
<td>2.80 (.70)</td>
</tr>
<tr>
<td>IN + QA</td>
<td>24</td>
<td>4.68 (1.19)</td>
<td>5.53 (1.12)</td>
<td>4.18 (.69)</td>
<td>2.75 (5.21)</td>
<td>5.08 (4.62)</td>
<td>2.25 (3.72)</td>
<td>3.61 (.55)</td>
<td>2.95 (.68)</td>
</tr>
<tr>
<td>IN + QA + Outline</td>
<td>28</td>
<td>4.87 (.96)</td>
<td>5.35 (1.15)</td>
<td>4.29 (.78)</td>
<td>.96 (4.29)</td>
<td>4.36 (4.56)</td>
<td>2.43 (2.92)</td>
<td>3.70 (.62)</td>
<td>2.71 (.82)</td>
</tr>
<tr>
<td>IN + QA + Outline + quizzes</td>
<td>30</td>
<td>4.74 (.87)</td>
<td>5.51 (1.17)</td>
<td>4.22 (.80)</td>
<td>.67 (4.42)</td>
<td>2.27 (6.16)</td>
<td>.53 (3.55)</td>
<td>3.71 (.65)</td>
<td>2.67 (.78)</td>
</tr>
<tr>
<td>Real Mentor</td>
<td>25</td>
<td>4.75 (1.25)</td>
<td>5.43 (1.27)</td>
<td>4.08 (1.04)</td>
<td>1.68 (5.09)</td>
<td>2.60 (4.55)</td>
<td>1.16 (4.07)</td>
<td>3.47 (.58)</td>
<td>2.94 (.86)</td>
</tr>
<tr>
<td>Total</td>
<td>158</td>
<td>4.69 (1.04)</td>
<td>5.39 (1.12)</td>
<td>4.10 (.85)</td>
<td>1.56 (4.73)</td>
<td>3.03 (5.34)</td>
<td>1.28 (3.78)</td>
<td>3.58 (.60)</td>
<td>2.81 (.79)</td>
</tr>
</tbody>
</table>

Table 12. Regression Coefficients of Activity

<table>
<thead>
<tr>
<th>Model</th>
<th>B</th>
<th>Std. Error</th>
<th>Beta</th>
<th>Correlation</th>
<th>t</th>
<th>Sig.</th>
<th>Zero-order</th>
<th>Partial</th>
<th>Part</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>21.609</td>
<td>10.360</td>
<td></td>
<td></td>
<td>2.086</td>
<td>.040</td>
<td>-0.094</td>
<td>-0.101</td>
<td>-0.096</td>
</tr>
<tr>
<td>STRATEGY</td>
<td>-2.217</td>
<td>2.216</td>
<td>-1.43</td>
<td>-1.001</td>
<td>-319</td>
<td>.004</td>
<td>-1.015</td>
<td>-0.047</td>
<td>-0.045</td>
</tr>
<tr>
<td>INTRING</td>
<td>-681</td>
<td>1.451</td>
<td>-0.55</td>
<td>-0.469</td>
<td>-640</td>
<td>.115</td>
<td>-0.028</td>
<td>-0.109</td>
<td>-0.104</td>
</tr>
<tr>
<td>EXTRING</td>
<td>-1.331</td>
<td>1.223</td>
<td>-1.20</td>
<td>-1.088</td>
<td>-279</td>
<td>.002</td>
<td>-0.225</td>
<td>-0.114</td>
<td>-0.109</td>
</tr>
<tr>
<td>ACT</td>
<td>.598</td>
<td>.262</td>
<td>.229</td>
<td>2.286</td>
<td>.024</td>
<td>.198</td>
<td>.114</td>
<td>.148</td>
<td>.148</td>
</tr>
<tr>
<td>SNS</td>
<td>9.404E-02</td>
<td>.235</td>
<td>.041</td>
<td>.400</td>
<td>.690</td>
<td>.065</td>
<td>.040</td>
<td>.038</td>
<td>.038</td>
</tr>
<tr>
<td>SEQ</td>
<td>.372</td>
<td>.327</td>
<td>.119</td>
<td>1.139</td>
<td>.257</td>
<td>.129</td>
<td>.114</td>
<td>.109</td>
<td>.109</td>
</tr>
<tr>
<td>DEEP</td>
<td>.521</td>
<td>.338</td>
<td>.204</td>
<td>1.542</td>
<td>.126</td>
<td>.003</td>
<td>.154</td>
<td>.148</td>
<td>.148</td>
</tr>
<tr>
<td>SURFACE</td>
<td>.216</td>
<td>.211</td>
<td>.106</td>
<td>1.025</td>
<td>.308</td>
<td>.114</td>
<td>.103</td>
<td>.098</td>
<td>.098</td>
</tr>
</tbody>
</table>

a. Dependent Variable: ACTIV_2

The second regression analysis had the eight learning motivation/strategy/styles measures as independent variables and the learning performance gain scores (in session 1, in session 2, and overall) as the dependent variables. Since we were most interested in the relationships between these variables for students learning with virtual mentors, we
conducted this regression analysis only to the virtual mentor groups (group 1 to 5). Results showed that the overall learning performance gain scores were significantly regressed on two measures: SNS, $\beta = -0.188$, $t(107) = -2.020$, $p = .046$; ExtrinsicG, $\beta = -0.262$, $t(107) = -2.687$, $p = .008$ (see Table 13). Therefore hypothesis 6 was partially supported for the overall learning process. Students who were less extrinsic goal driven and students who were more intuitive types of learners (lower score on SNS) would have more overall learning gain when learning with virtual mentors.

Table 13. Regression Coefficients for Overall Learning Performance Gain

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td></td>
<td>Zero-order</td>
</tr>
<tr>
<td>1</td>
<td>(Constant)</td>
<td>.933</td>
<td>.230</td>
<td>4.065</td>
<td>.000</td>
</tr>
<tr>
<td>DEEP</td>
<td>4.607E-03</td>
<td>.007</td>
<td>.074</td>
<td>.659</td>
<td>.511</td>
</tr>
<tr>
<td>SURFACE</td>
<td>-2.22E-03</td>
<td>.004</td>
<td>-.046</td>
<td>-.494</td>
<td>.622</td>
</tr>
<tr>
<td>ACT</td>
<td>-7.96E-05</td>
<td>.006</td>
<td>-.001</td>
<td>-.014</td>
<td>.989</td>
</tr>
<tr>
<td>SNS</td>
<td>-1.01E-02</td>
<td>.005</td>
<td>-.188</td>
<td>-.202</td>
<td>.046</td>
</tr>
<tr>
<td>SEQ</td>
<td>2.549E-03</td>
<td>.007</td>
<td>.033</td>
<td>.356</td>
<td>.723</td>
</tr>
<tr>
<td>STRATEGY</td>
<td>4.247E-02</td>
<td>.046</td>
<td>.116</td>
<td>.915</td>
<td>.362</td>
</tr>
<tr>
<td>INTRING</td>
<td>-4.42E-02</td>
<td>.031</td>
<td>-.148</td>
<td>-.200</td>
<td>.152</td>
</tr>
<tr>
<td>EXTRING</td>
<td>-6.96E-02</td>
<td>.026</td>
<td>-.262</td>
<td>-.687</td>
<td>.008</td>
</tr>
</tbody>
</table>

a. Dependent Variable: ARCALL
b. Selecting only cases for which TREAT ID IN ORDER < 6.00

Table 14. Regression Coefficient for Learning Performance Gain in Session 2

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td></td>
<td>Zero-order</td>
</tr>
<tr>
<td>1</td>
<td>(Constant)</td>
<td>.745</td>
<td>.181</td>
<td>4.110</td>
<td>.000</td>
</tr>
<tr>
<td>DEEP</td>
<td>5.627E-03</td>
<td>.006</td>
<td>.116</td>
<td>1.020</td>
<td>.310</td>
</tr>
<tr>
<td>SURFACE</td>
<td>-6.14E-03</td>
<td>.004</td>
<td>-.162</td>
<td>-1.728</td>
<td>.087</td>
</tr>
<tr>
<td>ACT</td>
<td>-2.34E-03</td>
<td>.004</td>
<td>-.048</td>
<td>-.529</td>
<td>.597</td>
</tr>
<tr>
<td>SNS</td>
<td>-6.02E-04</td>
<td>.004</td>
<td>-.014</td>
<td>-.152</td>
<td>.880</td>
</tr>
<tr>
<td>SEQ</td>
<td>1.121E-02</td>
<td>.006</td>
<td>.183</td>
<td>1.980</td>
<td>.050</td>
</tr>
<tr>
<td>STRATEGY</td>
<td>-1.44E-02</td>
<td>.037</td>
<td>-.050</td>
<td>-.393</td>
<td>.695</td>
</tr>
<tr>
<td>INTRING</td>
<td>-3.11E-02</td>
<td>.024</td>
<td>-.133</td>
<td>-.285</td>
<td>.201</td>
</tr>
<tr>
<td>EXTRING</td>
<td>-3.93E-02</td>
<td>.020</td>
<td>-.189</td>
<td>-.1923</td>
<td>.057</td>
</tr>
</tbody>
</table>

a. Dependent Variable: ARCMP2
b. Selecting only cases for which TREAT ID IN ORDER < 6.00
The learning performance gain score in session 1 was not significantly regressed on any of the eight measures, but in session 2 it was significantly regressed on one of the learning style measures: SEQ, $\beta = .183$, $t(107) = 1.980$, $p = .050$ (see Table 14). However, it was opposite to our expectations. Considering that the SEQ scale had very low reliability score in this study (Cronbach Alpha = .38), this result was ignored.

A correlation analysis was conducted to test the relationships among the learning motivation/strategy/styles measures and three students’ self-report measures (perceived learning effectiveness, perceived interactivity, and learning system preference) in each session. Again, this analysis was conducted only on the virtual mentor groups (groups 1 to 5). We found that in session 1, the following variables were correlated: DEEP and INTER, $r(131) = .183$, $p = .035$; ACT and INTER, $r(131) = -.280$, $p = .001$; ACT and LSP, $r(131) = -.249$, $p = .004$. In session 2, the following variables were correlated: SURFACE and LSP, $r(124) = .175$, $p = .049$; ACT and EFFECT, $r(124) = -.217$, $p = .015$; ACT and INTER, $r(124) = -.279$, $p = .002$; ACT and LSP, $r(124) = -.222$, $p = .012$. All these results were opposite what we expected in hypotheses 5 to 8. Particularly the ACT variable showed strong relationships to students’ perceptions in both sessions. It seemed that although more active type of learners did have more interactive activities in the advanced session, they had less perceived learning effectiveness, interactivity, and preference on learning systems. One possible explanation to these unexpected results is that active learners had more expectations on interactions than what the virtual mentor could provide. This discrepancy generated their dissatisfaction towards the virtual mentor. The qualitative results later provided some support for this finding.
Finally, we did not find any significant correlation among the actual interactive learning activities, the learning performance, perceived learning effectiveness, self-report interactivity, and learning system preference. Hypothesis 9 and 10 were not supported.

### 6.2.2.3 User Perception to System Features

Besides the four scales about students’ perceived learning effectiveness, interactivity, system usability, and learning system preference, there were twelve more five-point Likert-type questions in the post-experiment questionnaires asking students about their perceptions on some specific system features. These questions were included to see if there were any other possible factors influencing the relationship among constructs. Because not all features were available in every group, for some questions we added a sixth response choice of “not applicable”. Table 15 presents the details of the responses to these items in both session 1 and session 2. Questions asked about similar features are presented in the same block (the real questionnaire had these questions in a random order). Response frequencies, the standard deviation, the mean and the median for each question are listed. Because of the highly exploratory nature of these questions, only these descriptive statistics were used in interpreting the data.
Table 15. Responses to Questions about System Features

<table>
<thead>
<tr>
<th>Questions</th>
<th>Session</th>
<th>Strongly Disagree (1)</th>
<th>Disagree (2)</th>
<th>Neutral (3)</th>
<th>Agree (4)</th>
<th>Strongly Agree (5)</th>
<th>Missing or N/A</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. I enjoyed the self-paced control I had of the learning process. (I like being able to select and view any part of the lecture at any time.).</td>
<td></td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>12</td>
<td>75</td>
<td>39</td>
<td>1</td>
<td>4.10</td>
<td>80.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>14</td>
<td>64</td>
<td>40</td>
<td>0</td>
<td>4.06</td>
<td>90.0</td>
</tr>
<tr>
<td>B. The structured and synchronized multimedia content (the synchronously displayed video, slides and notes of lecture segments) in the system helped me understand the subject matter.</td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>18</td>
<td>83</td>
<td>31</td>
<td>1</td>
<td>4.10</td>
<td>60.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>31</td>
<td>59</td>
<td>25</td>
<td>2</td>
<td>3.78</td>
<td>91.0</td>
</tr>
<tr>
<td>C. There was too much information (video, slides and notes) on the screen at one time, making it hard for me to concentrate.</td>
<td></td>
<td>1</td>
<td>8</td>
<td>20</td>
<td>26</td>
<td>67</td>
<td>11</td>
<td>1</td>
<td>3.40</td>
<td>1.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>6</td>
<td>14</td>
<td>20</td>
<td>67</td>
<td>18</td>
<td>1</td>
<td>3.62</td>
<td>1.02</td>
</tr>
<tr>
<td>D.1. The &quot;Ask a Question&quot; function helps me find the specific information easily and quickly without going through the whole video.</td>
<td></td>
<td>1</td>
<td>4</td>
<td>8</td>
<td>30</td>
<td>34</td>
<td>15</td>
<td>42</td>
<td>3.53</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>3</td>
<td>8</td>
<td>14</td>
<td>32</td>
<td>29</td>
<td>40</td>
<td>3.88</td>
<td>1.09</td>
</tr>
<tr>
<td>D.2. The lecture outline which has links to specific topics helps me find the specific information easily and quickly without going through the whole video.</td>
<td></td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>15</td>
<td>61</td>
<td>32</td>
<td>24</td>
<td>4.14</td>
<td>67.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>8</td>
<td>10</td>
<td>66</td>
<td>23</td>
<td>18</td>
<td>19</td>
<td>3.97</td>
<td>78.0</td>
</tr>
<tr>
<td>D.3. The &quot;Ask a Question&quot; function helps me better on finding specific information than the lecture outline which has links to specific topics.</td>
<td></td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>34</td>
<td>30</td>
<td>12</td>
<td>44</td>
<td>3.45</td>
<td>.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>3</td>
<td>11</td>
<td>13</td>
<td>35</td>
<td>22</td>
<td>42</td>
<td>3.74</td>
<td>1.10</td>
</tr>
<tr>
<td>D.4. The Pop-up Questions helps me on learning/memorizing/reinforcing concepts.</td>
<td></td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>18</td>
<td>22</td>
<td>12</td>
<td>77</td>
<td>3.73</td>
<td>.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>28</td>
<td>21</td>
<td>3</td>
<td>68</td>
<td>3.33</td>
<td>.82</td>
</tr>
<tr>
<td>D.5. The questioning and answering process (including the &quot;Ask a Question&quot; function and/or the Pop-up questions) helped me better understand the concepts presented in the lecture.</td>
<td></td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>34</td>
<td>35</td>
<td>10</td>
<td>49</td>
<td>3.58</td>
<td>.81</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>4</td>
<td>7</td>
<td>19</td>
<td>38</td>
<td>9</td>
<td>49</td>
<td>3.53</td>
<td>.99</td>
</tr>
<tr>
<td>E.1. The video of the instructor in the lecture made me feel more involved in my learning process.</td>
<td></td>
<td>1</td>
<td>7</td>
<td>25</td>
<td>38</td>
<td>52</td>
<td>11</td>
<td>0</td>
<td>3.26</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>8</td>
<td>30</td>
<td>45</td>
<td>36</td>
<td>6</td>
<td>1</td>
<td>3.02</td>
<td>.99</td>
</tr>
<tr>
<td>E.2. The video of the instructor in the lecture didn't provide help in my learning process.</td>
<td></td>
<td>1</td>
<td>5</td>
<td>23</td>
<td>30</td>
<td>55</td>
<td>20</td>
<td>0</td>
<td>3.47</td>
<td>1.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>10</td>
<td>23</td>
<td>28</td>
<td>53</td>
<td>12</td>
<td>0</td>
<td>3.27</td>
<td>1.11</td>
</tr>
<tr>
<td>E.3. The video quality of the lecture was satisfying.</td>
<td></td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>23</td>
<td>84</td>
<td>20</td>
<td>1</td>
<td>3.88</td>
<td>.76</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>24</td>
<td>77</td>
<td>12</td>
<td>0</td>
<td>3.67</td>
<td>.87</td>
</tr>
<tr>
<td>E.4. The audio quality of the lecture was satisfying.</td>
<td></td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>11</td>
<td>80</td>
<td>34</td>
<td>0</td>
<td>4.05</td>
<td>.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>3</td>
<td>9</td>
<td>23</td>
<td>73</td>
<td>17</td>
<td>1</td>
<td>3.74</td>
<td>.87</td>
</tr>
</tbody>
</table>
Two system features other than interactivity received very positive responses. Students really liked the self-paced control in their learning process in both sessions (question A). Most of the students agreed that the structured and synchronized multimedia were very effective in help them learn (question B), especially in the first session. However, we did find from question C that some students experienced an information overload problem with this multimedia representation, especially in the second session. Therefore, it seems that fewer, but personalized media channels are more suitable for the advanced learning session.

Question D.1 to D.5 asked students directly about their experience with the QA, outline, and pop-up quizzes functions. An interesting finding from these items is that students’ satisfaction towards the outline decreased from session 1 to session 2, while their satisfaction towards QA increased from session 1 to session 2. In addition, in session 2 more students thought QA was better than outline in finding specific information. Therefore, it seems that the 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.

Finally, from responses to question E.1 to E.4 we saw that there was neutral opinion on whether the video of the instructor facilitated learning. We will discuss this issue based on students’ responses to the open-ended questions in the next section.
6.2.2.4 Open-ended Responses

In addition to the quantitative data collected, the qualitative data collected by the open-ended items in the post-experiment questionnaires provided supplemental findings and feedbacks for in-depth understanding of the research problem and further development efforts.

There were eight open-ended questions in the post-experiment questionnaire for virtual mentor groups in each session. Four out of these eight questions that were not directly about the system features were modified and asked in the real mentor group. The actual questions and their responses were listed and analyzed later in this section.

Responses to these open-ended questions were analyzed by theme using content analysis, in which verbal content was coded using a quasi-objective scheme, and then further analyzed using relative frequency statistics. A software program *CDC EZ-Text* was used to support the coding and analysis of these open-ended responses (Carey et al. 1998). The author created a codebook in which each code represented a unique theme in the responses. The codebook was developed incrementally as the author proceeded through the responses. The final codebook contained codes for a total of 141 discrete themes (ideas) for all responses in both sessions (see Appendix A). Each individual response to a question had at least one theme but might contain different themes, and therefore might be assigned several different codes. The same code could also be assigned to responses to different questions. After all responses were assigned with codes, the coding results were exported into an EXCEL spreadsheet for further frequency analysis.
The actual questions and their responses are discussed below. For each question, we listed the most common themes (code frequency >=5) and their frequencies in a table. The implication of these themes is discussed; with support from direct quotes of student responses (no typos were corrected). These qualitative results are also discussed in conjunction with the quantitative results described in the previous sections.

**Question No. 1: What do you like about the system? (“What do you like about the lecture?” in the Real Mentor group)**

This question asked the students to list the factors that determined their satisfaction towards the system or the lecture. Although we were mostly interested in interaction in learning, we left this question very open so that we could understand the role of interaction in students’ learning process.

<table>
<thead>
<tr>
<th>Session</th>
<th>N</th>
<th>3PANEL</th>
<th>EASYUSE</th>
<th>CONTENT</th>
<th>LOTINFO</th>
<th>OWNPACE</th>
<th>QA</th>
<th>REVIEW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group1</td>
<td>1</td>
<td>26</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>22</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Group2</td>
<td>1</td>
<td>25</td>
<td>7</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>24</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>Group3</td>
<td>1</td>
<td>24</td>
<td>5</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>23</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Group4</td>
<td>1</td>
<td>28</td>
<td>10</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>25</td>
<td>6</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Group5</td>
<td>1</td>
<td>30</td>
<td>4</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>29</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td>3</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>Group6</td>
<td>1</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The most common themes in the responses to this question were listed for each group in each session in Table 16. Despite that we thought interaction, especially QA, is very important for students’ learning, most students liked the system just because of the self-
paced control provided in all system groups (OWNPACE). Because the lecture was pre-
recorded, they could pause, skip, or replay any part of the lecture video, without the 
limitation of time or location. Even for the IN only group where the students had least 
control over the lecture (only standard pause, fast forward, and rewind buttons were 
available), they liked the system because of this control. At the same time, students also 
mentioned that they liked the convenience of reviewing specific content in the system 
(REVIEW). There was not much difference in the amount of these two themes between 
sessions or among groups. Typical comments include:

they help me learn at my own pace. I usually have problems in class because I 
can’t take notes fast enough or the teacher talks to fast for me to understand 
things before moving on to another topic. [#314 in session 1]

I like that you can do the class at your own pace and keep going back to the 
lecture. It is not a one time thing like in the traditional setting. [#208 in 
session 1]

I like being able to move through the lecture at my own pace. In a classroom 
you don’t get to ask the teacher to repeat as much as may be necessary. This 
program allows for the student to actually get the answer they need. [#521 in 
session 2]

Another two common themes across different groups and sessions were that students 
liked 1) the structured and synchronized multimedia lecture that integrated the lecture 
video, slides and notes (3PANEL); 2) the ease of use of the system (EASYUSE). For 
example, students stated that:

Having videos and powerpoint slides next to one another on the screen makes 
it seem like attending a real lecture. Also, the outline for the presentation was 
helpful in keeping me focused on exactly what I was supposed to be learning. 
[#123 in session 1]
I liked that there was different ways that you could learn the material. The notes on the bottom of the screen helped keep my attention on the lecture rather than off in lala land. [#202 in session 1]

I liked that it was all synchronized as far as the text, powerpoint, and video. There was also a lot of information that was easy to take in in a relatively short amount of time. [#322 in session 1]

Ease of use, easy to skip information already acquired easy to go back to something you may have missed and review. [#517 in session 1]

I thought the navigation was very easy and made it much better to learn with than a traditional system where you would watch the whole thing and then have to find the place you wanted to watch again [#408 in session 2]

The last four themes were consistent with the responses to the closed-ended questions where all these items had means higher than 4.0. This also explained to some extent why the five virtual mentor groups did not have much difference among learning outcomes.

The most important feature of e-learning to students, in this study, was not the interaction but the self-paced control and other issues that were available in each group. The interaction might still be important, but not as important as these common features.

The only common theme that was different among groups and between sessions was about QA. When QA functions were available (group 3, 4, and 5), students did include QA as one of the reasons why they liked the system; while only a few students included the outline or the pop-up quiz as one of the reasons. In addition, there were twice or three times more students in the second session who listed QA as one of the reasons they liked the system than in the first session. This again confirmed that interaction especially QA is more important for the advanced learning phase.

…I like the independence I have in learning the material without being in a classroom full of students that I do not care about and have no baring on my learning curve. The program was very easy to access and probe. I love the
Ask the Question box for specific words or terminology that I can easily get to. … [#417 in session 1]

very easy to use. very nicely done. i love the ask me a question part. i liked seeing the teacher and his slides to his speech. the questions in the middle were very helpful [#518 in session 1]

I like how you can go at your own pace and ask questions I might not feel comfortable asking in class. [#310 in session 1]

I enjoyed that you could jump from one part of the lecture to the next. I also liked that you could either type in key word or sentences, if you please, into the ask a question box [#403 in session 2]

Finally, when asked in the Real Mentor group, the most common response was about the lecture content. Therefore, when learning with a real mentor, the students cared more about the content than other issues such as interaction. They would more possibly like the lecture when they thought the information was important for them to learn.

It was useful information, that could be helpful to almost anyone who uses computers. [#605 in session 1]

I did find many important fact and information that I hope to use to protect myself from being accessible. [#601 in session 1]

My knowledge of computer security before the workshop was very low, know I feel like I know a little bit about what is going on. [#605 in session 2]

Question No. 2: What do you dislike about the system? (“What do you dislike about the lecture?” in the Real Mentor group)

This question asked the students to list the factors that determined their dissatisfaction towards the system or the lecture. The responses for this question were more diverse than the last question and this time interaction played an important role of in students’ learning process.
The most common themes in the responses to this question were listed for each group in each session in Table 17.

One thing we did not expect before the experiment was students’ reaction to the instructor’s image in the lecture video. Some students did not like the system or the lecture just because they thought the instructor was annoying (ANNOY) and the lecture was boring (BORE). For example,

*The instructors constant stumbling over words, the inconstistency in what was being said and what was written in the text box, the instructors figiting around. [#414 in session 1]*

*The virtual instructor was very very monotone and kept coughing in between every other sentence. He was reading from the screen a lot and did not seem excited or interested in what he was lecturing about, and this in turn turned me off to the lecture. It was very easy to get distracted and was hard to concetrate on him for long periods of time because he was too monotone and boring. [#203 in session 1]*

*Watching the computer screen is boring for me. In a classroom I can focus on the teacher moving around. [#505 in session 1]*

*It was very boring and hard to pay attention , the instructor seemed like he was just reading a script and not explaing. [#524 in session 2]*

<table>
<thead>
<tr>
<th>Session</th>
<th>N</th>
<th>ANNOY</th>
<th>BORE</th>
<th>NOFOCUS</th>
<th>OVERLOAD</th>
<th>OUTLINE</th>
<th>NOQA</th>
<th>QALIMIT</th>
<th>PERSON</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>26</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>9</td>
<td>6</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>22</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>11</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Group2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>25</td>
<td>4</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>24</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Group3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>23</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>23</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Group4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>28</td>
<td>5</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>28</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Group5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>30</td>
<td>3</td>
<td>8</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>29</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Group6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>22</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>21</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
In addition, some students who took the class from the same instructor before found that the instructor acted differently in the lecture video.

XXX’s script is boring. He usually is more animated in person and has more fascinating stories during IN class sessions that what he presented in this experiment. I think a little bit of humor and real life experiences or examples could make the speech a little more interesting. But then again, too much of it can be over bearing and time consuming [#417 in session 1]

It seemed that the instructor unintentionally changed his behavior when being videotaped for the e-learning system. Considering that similar things happened with the Agent99 Trainer study (users disliked Agent99 Trainer because the instructors appeared to be fidgeting in the lecture videos), this phenomenon might not be a single chance event. Different instructors and their different behaviors might significantly affect students’ learning outcomes. Therefore, it is necessary that we introduce a better way to control the behavior of instructors in the future.

Another two common complaints about the system were that 1) there was too much information in the system (OVERLOAD); and 2) students could not focus on learning (NOFOCUS). Although many students liked the combination and synchronization of video, slides and notes, some students faced the information overload problem and had difficulty to really concentrate in learning. Students also found that they could not focus on learning because of no obligation/attention from the virtual mentor. This confirmed our finding that students who were more extrinsic goal oriented would not learn effectively with the virtual mentor. They have to be self-disciplined to be successful in learning with virtual mentors. The first session seemed to have more these two complaints than the second session, as shown in Table 17. This was probably because
they were getting used to the system after the first session. We quoted some comments below.

While it was good that all of the info was on the same screen, it was hard to concentrate on the instructor because there was too much information to take in. I found myself missing some of what the instructor was saying because I was trying to read the outline to get a general idea of what he was trying to say. [#430 in session 1]

you don’t feel bad dozing off infront of a computer program... in a class room you have to atleast make the effort to LOOK interested, which sometimes helps you pay attention [#224 in session 1]

I think that it allows me to day draem more and that if i was at home watching this there is no way i could stay focused with out someone actually teaching me in a classroom. [#503 in session 2]

I dislike the fact that it is so easy to get distracted. I found myself not always concentrating on the video because I would be looking at the notes or reading the transcript. In the same way, I also focused so much on the video at times that I forgot to view the outlines or the transcript. [#213 in session 2]

Lack of interaction was indeed counted as an important reason for students disliking the system. In group 1 and group 2 where the QA function was not available, many students complained that they could not ask questions to the instructor (NOQA). They said that:

I did not like that you could not ask question about a certain topic if you had trouble with it and you could not go back in the lecture to try to understand it better [#121 in session 1]

There is no way to ask questions. It does not hold your attention because the lector has no way to respond to the class. [#221 in session 1]

It’s incredibly boring; there is not any interaction with the instructor/classmates; no questions are asked and no answers are given. [#214 in session 2]

In group 1 and group 3 where the hyperlinks on the outline were removed, many students complained that they could not quickly navigate through the lecture (OUTLINE).
Students in group 1 were extremely unhappy with the fact that they had to use the fast-forward/rewind buttons. Therefore, although adding the outline hyperlinks might not make the students like the system more, removing them could result in inconvenience in learning and thus negative feelings towards the system. Comments include:

*I don’t like how you can’t select a topic form the outline and go straight to that slide i without having to first ask a question. [#310 in session 1]*

*It is difficult to navigate through. I wanted to be able to use the outline to click my way to various topics but I had to use fast forward. When I fast forwarded it would stop at each topic in the lecture and pause to load information. If I wanted to go from the beginning to the end, it would take a very long time. [#107 in session 2]*

*I could not click on the subject topics in the outline. I found this very annoying and discouraging because it took me a lot longer to use the "Ask a question feature" when I did not have a specific question such as defining something. When I needed to compare two items or find a broad concept, I could not just graze through the text. I had to "fast forward" through the video until I got to the topic. This took a long time and frustrated me because I didn’t complete the assignment. [#326 in session 2]*

When the QA function was available (group 3, 4, and 5), although students did include QA as one of the reasons why they liked the system, they found that the QA function was not as good as they imagined (QALIMIT). Because the QA function was completely based on the transcribed speech in the lecture video, students’ questions might not be answered if they were outside of the lecture scope. In addition, it was kind of impersonal and uncomfortable for students to ask questions to a computer other than talk with a live instructor or classmates (PERSON). For example,

*when i asked a couple of questions, it was not able to understand my questions. i think the words that i type in need to be more broad to bring up all the different areas of the discussion. [#421 in session 1]*
I did not like that you cannot ask questions that are not included in the lecture. There should be some sort of extra information for the students who want to learn more about a topic. [#423 in session 2]

certain questions might not get answered or might not get answered in the way i would want them to. the questions asked only refer to the info in the slides but if i had a question outside of that straightforward info the program might not be as useful or helpful [#302 in session 2]

It was not very interactive with others. I do better when I can interact with more than one person. They can understand abstract questions better as well. [#506 in session 1]

It is not the same as being in the class. I feel like I dont learn because there is no involvement. No eye contact or interaction sucks. [#416 in session 2]

It takes away from group interaction, and professor student interaction. Questions were not answered the way wanted. [#519 in session 2]

Finally, when asked in the Real Mentor group, most students said that there was nothing that they disliked (6 out of 22 responses in session 1 and 7 out of 21 responses in session 2). Therefore, although based on our observation most students did not directly interact with the instructor, they still liked learning with a live instructor more than with a computer.

Question No. 3: How can the system be improved to better help you learn?

Responses to this question were somewhat similar to those for the question No. 2, since it asked students how to solve the problems they found with the system.

The most common themes in the responses to this question were listed for each group in each session in Table 18.
First of all, some students suggested changing the instructor (ANNOY).

*Have a more enthusiastic instructor because I was getting bored and did not want to listen to him anymore. [#405 in session 1]*

*The instructor would be more interesting if he looked at me as if he were really speaking to me. [#410 in session 2]*

Some students also thought of ways to change the appearance of the user interface (APPEAR) to solve the information overload problem.

*have the lecture scroll automatically [204]*

*highlight the main terminology because it is easy to find the word. [206]*

*Bigger instructor screen [222]*

*Have to option to close windows that you do not want to look at. It gets hard to look at all of the information on the screen. Sometimes, I can’t decide what to look at. [318]*

Again, groups with different levels of interactiveness asked for different ways to improve the system. For example, better navigation and/or interaction (MOREINT/NAVIGAT), especially providing QA in group 1 and group 2 (QA); adding links on the outline in group 1 and group 3 (OUTLINE); and providing better, expanded
answers in group 4 and 5 when both QA and hyperlinked outline were available (EXPAND).

I like the opportunity to ask questions and have a discussion, if there was a way to make it more interactive and personal that would be better, and maybe not include the notes from the speaker so that it would make it easier to take notes and read the slides. [#114 in session 1]

A speaker with some enthusiasm would help tremedously. And also have a place where you can ask questions if you want. It would help if the learning was interactive, either you could chat with the lecturer right there, or there were tests you could give yourself right after hearing the lecture, so see if you understood and learned everything that you needed to [#203 in session 1]

Make it more hands on and allow me to ask questions and have the instructor ask me questions and to say this is important or this isn't important. [#124 in session 2]

I think that there needed to be more interaction between the user and the lecturer. If I had a question on the material it would not be answered because there was no one to ask. Even if it went to an email address to be answered would satisfy me. [#202 in session 2]

Direct links to specific topics, more interaction with the Instructor. [#105 in session 1]

The video needs to be larger, I think you can get rid of the transcript and slide number, I did not look at it. You could add a function where you click on the slide number and it takes you directly to that portion of the lecture. [#208 in session 1]

I hope that I can jump to the other lecture of the outline. Pushing the passing key takes long time to got to other topic of the lecture. [#109 in session 2]

I think that the system could be improved if you could provide information that I might find interesting that was not directly covered in the lecture. [#423 in session 1]

Should include some kind of discussion area with the profesor and other students. When asking the question, I would like to see some other answer than just repeating lecture [#504 in session 1]

If the system had links to informational web sites it would help me learn. Also if it provided optional information to expand on a topic. [#403 in session 2]
Possibly have different explanations of each concept, so if the user didn’t understand a topic the first time, he could have it explained to him a different way. Possibly there could be something to click on that explains certain things in more depth, but if you understood the concept, you didn’t have to use that function and could continue on with the lecture. [#430 in session 2]

Finally, quite a few users in the full-function group (group 5) viewed the system as a very well-designed system and responded with “n/a” or “none” (NONE), or comments such as:

*I thought it was excellent. I have no complaints. [#520 in session 1]*

*The system is good as is, I think that it only tailors to a select group of learners that prefer a virtual instructor over a real living teacher. [#503 in session 2]*

**Question No. 4: In what learning situations do you think you would like to use this type of learning system?**

The responses to this hypothetical question were also very diverse. Only the six themes listed in Table 19 could be called “common” for certain groups.

<table>
<thead>
<tr>
<th>Group</th>
<th>Session</th>
<th>N</th>
<th>ANY</th>
<th>ENTRY</th>
<th>FACT</th>
<th>LECTURE</th>
<th>ONLINE</th>
<th>SPEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group1</td>
<td>1</td>
<td>25</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>22</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Group2</td>
<td>1</td>
<td>24</td>
<td>6</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>24</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Group3</td>
<td>1</td>
<td>24</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>21</td>
<td>1</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Group4</td>
<td>1</td>
<td>27</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>28</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Group5</td>
<td>1</td>
<td>30</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>29</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>
First, quite a few students, especially in group 1 and group 2 in session 1, and in group 5 thought that this system could be used in any or most situations (ANY), with comments like:

*I would like to use this system in every instance possible. With my learning style, I think that I would benefit greatly from it while relaxing at home at the same time.* [#118 in session 1]

*Any, but its probably best for non-hands-on learning.* [#226 in session 1]

*I would love to be able to have all lectures taped and presented like this so if after class you wanted to review what the speaker was talking about you could.* [#502 in session 1]

*In most of my classes, especially when the weather is bad or my kids are sick and I can’t make it to class.* [#505 in session 2]

*I would love to take classes this way! At my own pace and on my own time. I think this method is time effective. I wouldn’t have to waste my time trying to learn the traditional way. I could actually master the information instead of just doing whatever it takes to make the teacher happy to give you the grade you need.* [#521 in session 2]

Some students especially some in group 3 thought that only entry-level (ENTRY), fact-based (FACT) classes were suitable for this type of learning system. This might be because that the students found that the QA function were most effective for factoid questions. For more advanced questions like “how” type of questions, the current QA algorithm would not work well.

*In a class that requires a lot of factual knowledge as compared to concept based* [#303 in session 1]

*Definitely in IS courses such as this one. Also, any type of learning that requires you to attain a lot of detailed factual information such as history lessons or social sciences. Definitely not in situations requiring a lot of hands on activities such as math lessons and not in the arts such as communication courses either.* [#326 in session 1]
I think that in factual classes where theory isn’t an issue this format would be successful. Mathematical classes this method would not be successful. [#324 in session 2]

Entry level classes, rather than having a lecturer for 150 students in one classroom this would be a much more useful tool for learning [#302 in session 2]

Finally, several students in group 4 and 5 suggested that this system could be used in pure lecture-based classes where not much hands-on practice was needed. Or it could be used for classes that were completely delivered online. Some students also listed some specific subject matters that they thought would fit for this system such as Sociology, and some specific subject matters that they thought this system would not be useful such as Math.

I would have enjoyed it in my psychology class because it was all on-line and hears someone explain material is better then just reading it out of the book. [#403 in session 1]

I think it would be beneficial in most any class but mainly classes that have less hands on and more lecture. [#526 in session 2]

I would like to use this for classes that are based mostly on the lecture. It is nice to be able to repeat the parts of the lecture that I was a little fuzzy on or wanted review in. [#423 in session 2]

Question No. 5: Do you prefer this online learning system or traditional courses? Why?

We had hoped to see that there were more students who preferred the virtual mentor system in the groups with more interactivities. However, from the coded responses listed for each group in each session in Table 20, there was not obvious difference among groups, which was consistent with our previous quantitative results about Learning System Preference (LSP). For almost every group, there were definitely more students who preferred the traditional courses than the online learning system, although there were
also some students in each group who liked both of them. The reasons that the students provided for their preference also clustered well.

<table>
<thead>
<tr>
<th></th>
<th>Session N</th>
<th>SYSTEM</th>
<th>TRADIT</th>
<th>BOTH</th>
<th>PERSON</th>
<th>OWNPACE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group1</td>
<td>1 26</td>
<td>4</td>
<td>14</td>
<td>6</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2 22</td>
<td>4</td>
<td>13</td>
<td>2</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Group2</td>
<td>1 25</td>
<td>8</td>
<td>16</td>
<td>1</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2 24</td>
<td>6</td>
<td>14</td>
<td>3</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Group3</td>
<td>1 24</td>
<td>4</td>
<td>15</td>
<td>3</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2 23</td>
<td>6</td>
<td>15</td>
<td>2</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Group4</td>
<td>1 27</td>
<td>5</td>
<td>13</td>
<td>8</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2 28</td>
<td>4</td>
<td>16</td>
<td>4</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>Group5</td>
<td>1 30</td>
<td>8</td>
<td>17</td>
<td>5</td>
<td>14</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>2 29</td>
<td>6</td>
<td>18</td>
<td>5</td>
<td>11</td>
<td>4</td>
</tr>
</tbody>
</table>

Students preferred traditional courses over the LVM system mostly because they liked the personal interaction among students and the instructor. They thought the learning system removed all the non-verbal and social side of the interactions in learning. For example,

*No. Because I like going to class and having a real person talk to me, not a computer screen. [#103 in session 1]*

*No, I prefer traditional courses because I like the personal interaction between students and teacher...even if the teacher just lectures to 200 students for 50 minutes, that is still more personal and I feel more involved than watching someone on a video for 50 minutes [#203 in session 1]*

*I like traditional courses because I can pay attention better because you physically can interact, not sit there and then type a question you have. If the question was not answered then it might become frustrating not getting the exact answer you want. [#301 in session 1]*

*Other than what I mentioned above, I wouldn't want to use this type of learning system. I enjoy seeing facial features and feeling the grandness of someone's presence in real life, it just feels different. The excitement a teacher can bring and how the teacher grabs an entire classes' attention is far better. And spontaneous questions can be asked, and this way other students might generate*
questions at the same time, this interaction is not available in this system. I am more awake when I have someone talking to me or when a class is actively participating. I didn't feel the need to pay attention with this system. So, I prefer 9 out of 10 to attend a traditional class. [#412 in session 1]

tradition courses becasue i think it complete takes away from the social part of college. college is the best 4 years of your life, so attending it is all the fun. i also like studing in groups, meeting new friends through college. with the online system it elimates the whole social aspect of the college life. [#518 in session 1]

I prefer traditional courses because I feel more personally involved in my learning process. A traditional course gives me more of a real feeling that I am learning something, where this system was not capable of doing that. [#213 in session 2]

'I like the tradtional course becuase you can interact with the instructor and not sit there and not be able to talk back and verbally and nonverbally communicate with the instructor. [#301 in session 2]

No. Interaction, Non-verbal contact, and more are SO IMPORTANT in learning. If people preferred not being at events in person, no one would go to live sports events, debates, parades, or more. [#416 in session 2]

I prefer traditional courses because students gain a learning relationship with their instructor where they can ask specific questions that are geared only towards them [#524 in session 2]

On the other hand, students who preferred the LVM system over traditional courses were mostly attracted by the self-paced control and convenience of the system. They still thought the interaction in the learning system was not as good as the interaction with a live person in a traditional course. However, the convenience of the system overweighed the insufficient interactions. These students also usually reported that they were either shy and did not like to interact with other people, or had constraints in attending courses regularly (e.g., having busy work schedule or need to take care of a baby). Example comments were listed below.
I prefer online. Then I can learn conveniently from home. I am not a people-person, so avoiding people is always a bonus for me. [#226 in session 1]

Yes. Definitely. Because you're not surrounded by strangers and you are in control of your time and how much you want to learn. This program can not help you fail the class. It only helps you improve your grade and learning abilities. If you fail in such program, it's because you did not put any effort in it. Some people may think this is too easy but for me I learn a little slower than the rest. I love it. [#417 in session 1]

I thought that this system was creative because it gave you options to take breaks to think about and go over certain subjects. I thought that this gave a person more room for questions without disrupting the teacher or the class. [#527 in session 1]

YESS.. you can do it on your own time at your own pace and not feel like you sat through a useless lecture ever again, and you still feel the personal attention of a teacher. [#224 in session 2]

Online...you can ask it any question and it will give you a direct answer [#304 in session 2]

Yes. Because I like to self paced versus the traditional. Also, the video, script, and the other functions that help me ask questions, jump to different subjects help me understand and learn better and faster without being embarrassed asking a teacher during in class discussion. [#417 in session 2]

this online system because of my personal learning and lifestyle. I liked the integration of learning styles and I also have a child so it would be nice to do this at my own pace [#520 in session 2]

Finally, there were students who did not have specific preference between the LVM system and traditional courses. For most of these students, each choice had its own advantages in some specific situations, and the combination of these two would be the best way to learn.

I like both because online learning shows me exactly what is going on there and many details for me to understand the content. and traditional courses have conversation and reaction between me and instructor, and I can ask any question at that time. [#109 in session 1]
Traditional course for the availability of the teacher after class for specific questions. This online learning system so that I can review misunderstood issues at my leisure. [#414 in session 1]

yes and no, its easy to review and go over things, but the social aspect of being in a class and knowing both the students and professor can never and should never be replaced. Hybrid method doing both this online stuff and classes is probably a good way to go. [#517 in session 1]

a mixture of the two would be ideal. if i were given a choice of taking an online course or a traditional course, i would take the easier courses online and the more difficult classes in the traditional way. this way for the harder classes i would be able to go to tutoring or to the instructors office hours after class. [#118 in session 2]

Yes or No. This courses is good for students who study, but not good for students who don’t study, because no instructors are with them, those students will never participate in this course [#419 in session 2]

In beginner courses I like this better but in the more advanced smaller classes I would rather have an actual teacher. [#502 in session 2]

Question No. 6: When you need to find specific information in the lecture, how would you find it? Do you prefer asking questions or browsing the lecture outline? Why?

Our quantitative results showed that the QA function significantly increased the amount of students’ interactive activities in the second session. The outline function did not show that effect. Therefore, we would expect that there were more students who prefer asking questions than browsing the lecture outline in the second session. This was true for group 3 and group 4, but not for group 5. So we still could not confirm this expectation. We also found that the students’ preferences were definitely affected by their available system functions. For example, in group 1 and group 2 in which only the outline was available, most students said that they preferred browsing the outline.
Table 21. Codes Frequencies for Question 6 Responses

<table>
<thead>
<tr>
<th></th>
<th>Session</th>
<th>N</th>
<th>OUTLINE</th>
<th>QA</th>
<th>BOTH</th>
<th>EXACT</th>
<th>QUICK</th>
<th>EASYUSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group1</td>
<td>1</td>
<td>25</td>
<td>13</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>21</td>
<td>6</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Group2</td>
<td>1</td>
<td>24</td>
<td>13</td>
<td>9</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>24</td>
<td>17</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Group3</td>
<td>1</td>
<td>24</td>
<td>4</td>
<td>16</td>
<td>3</td>
<td>7</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>23</td>
<td>3</td>
<td>18</td>
<td>2</td>
<td>6</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Group4</td>
<td>1</td>
<td>28</td>
<td>12</td>
<td>12</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>28</td>
<td>7</td>
<td>14</td>
<td>7</td>
<td>4</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>Group5</td>
<td>1</td>
<td>29</td>
<td>16</td>
<td>10</td>
<td>3</td>
<td>8</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>29</td>
<td>15</td>
<td>7</td>
<td>6</td>
<td>9</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

The most common themes in the responses to this question were listed for each group in each session in Table 21. When students compared these two methods, there were three major criteria: whether it could find the answer exactly as needed (EXACT), the speed of finding the answer (QUICK), and how easy it was to find the answer (EASYUSE). It seemed that students who preferred QA mostly thought QA was faster and easier to find exact answers.

\textit{The lecture was very difficult to browse and the fast forward option was extremely ineffective. I prefer to ask questions because it gives me a better understanding of what the instructors point is.} [#1 13 in session 1]

\textit{Browsing the lecture outline was simpler, but asking the teacher for help will not only get that specific question answered the first way it was presented, but you have the chance of getting more information and having the teacher explain it in a different and more specific way.} [#203 in session 1]

\textit{Asking questions because it is faster and it finds similar topics that are related.} [#214 in session 1]

\textit{I like that. They answers as if they were our professors. When I found something I want to know, just typing my question, and they answer it even if there is professors there. That is pretty nice.} [#419 in session 1]
"ask a question" feature. I prefer asking questions. If I couldn’t figure out the question to ask, I browsed. Question got straight to it, browsing wastes time. [#215 in session 2]

Asking questions. Because 8 out of 10 I got an immediate answer from the ask function that was correct, and it provided another link for more info. or some topic related to [#412 in session 2]

This time, I asked specific questions, it helped me to get the right answer very quickly. [#504 in session 2]

On the other hand, students who preferred browsing mostly thought browsing was an easier thing to do especially when they had no specific questions.

I would probably browse the lecture outline because you can easily find the answer to your question in a few short moments. [#118 in session 1]

I would prefer browsing through a lecture outline, because I can go directly to the information I need. [#215 in session 1]

The browser and booknotes because I knew where it was in the lecture. A lecture with more detail might require the ask feature [#422 in session 1]

browsing the lecture outline after listening to the entire lecture. It was easier for me. I did not feel like I was asking the right questions to get the right response when I did ask questions. [#510 in session 1]

Browse the lecture outline. I can find it at my own pace and don’t have to be very specific w/the question to get the answer I want. [#207 in session 2]

I just browsed the outline because the headings were easy to find and were directly related to the questions asked [#508 in session 2]

For some other students, the combination of browsing and asking was simply the best way to find specific information in the system.

Both. sometimes I just didn’t catch something the first time. other times I just dont understand it the way it was explained [#110 in session 1]

I just double clicked on the section that I needed to review again. Also, I typed in a word that I was having problem understanding the concept. [#417 in session 1]
I used both the question and the outline. If I remembered exactly where the information I was looking for was in the lecture then I used the outline, but when I was looking for general information on a topic I used the ask feature.
[#423 in session 2]

It depends, if I just wanted to ask a quick question I asked, if I wanted more info I browsed [#520 in session 2]

Question No. 7: Do you like your interaction with the (virtual) instructor? Why? (“Do you like your interaction with the instructor? Why?” in the Real Mentor group)

Although we thought we provided virtual interactions to students in the LVM system, did the students really feel that they were interacting with the instructor? This question asked about the students’ subjective feeling about the interaction and compared it to the one as in the Real Mentor group. The results were encouraging. Although most students enjoyed their interaction with a real mentor, when more virtual interaction was provided, there were indeed more students who felt that they liked the interaction with the virtual mentor (LIKEINT) or at least the interaction was OK. When virtual interaction was not provided, such as in group1, most students thought there was no interaction (NOINT) at all. Although there were also many students who did not feel the interaction in the other groups, the number of these NOINT responses was decreased when more virtual interaction was provided. In addition, we found that several students did not like the interaction simply because the answers to their questions were the repetition of the lecture segments (REPINT). Therefore, it would be possible that these students would enjoy the virtual interaction in the future if expanded answers were provided.
The most common themes in the responses to this question were listed for each group in each session in Table 22. Several students wrote very positive comments about the virtual interaction.

Table 22. Codes Frequencies for Question 7 Responses

<table>
<thead>
<tr>
<th></th>
<th>Session</th>
<th>N</th>
<th>LIKEINT</th>
<th>OK</th>
<th>NOINT</th>
<th>REPINT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group1</td>
<td>1</td>
<td>24</td>
<td>1</td>
<td>0</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>21</td>
<td>2</td>
<td>3</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>Group2</td>
<td>1</td>
<td>25</td>
<td>4</td>
<td>0</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>24</td>
<td>1</td>
<td>2</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>Group3</td>
<td>1</td>
<td>24</td>
<td>4</td>
<td>6</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>23</td>
<td>4</td>
<td>7</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>Group4</td>
<td>1</td>
<td>28</td>
<td>11</td>
<td>4</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>28</td>
<td>5</td>
<td>6</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>Group5</td>
<td>1</td>
<td>28</td>
<td>9</td>
<td>4</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>29</td>
<td>5</td>
<td>5</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>Group6</td>
<td>1</td>
<td>22</td>
<td>13</td>
<td>1</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>22</td>
<td>13</td>
<td>1</td>
<td>9</td>
<td>0</td>
</tr>
</tbody>
</table>

yes because the instructor is same as my class instructor and he explain the detail deeply [#109 in session 1]

yes it gives a face to go with the voice [#323 in session 1]

Yes, of course. It is because I can really get the more correct and quick answer rather than serching the internet or asking classmates. [#406 in session 1]

The interaction was good, but not close enough to a real meeting with a professor. It was close enough though to real interactions for many students. [#408 in session 1]

Yes. He looks like real professor. [#501 in session 1 and 2]

yes, he did everything I ask him to do. [#513 in session 1]

yes, because it is almost identical with a live instructor. [#525 in session 1 and 2]

yes, because it could answer most of my questions, and very helpful [#518 in session 2]
yes because the instructor gives a feeling of him teaching you on a one-to-one basis. [#314 in session 2]

However, most students still thought there was no interaction in the system at all, even when virtual interaction was provided. One major reason for this response was very similar to why students prefer traditional courses – the personal side of the interaction.

For example,

*There wasn’t any interaction with the instructor. He basically read the notes and we couldn’t ask questions or even have him explain things more.* [#114 in session 1]

*No, you can’t ask specific questions, and he can’t answer them.* [#223 in session 1]

*I wouldn’t really call it interaction, it was just a guy on the computer screen talking. Interaction to me would mean if I could actually physically talk to someone.* [#302 in session 1]

*I didn’t feel like there was an interaction. I would rather just read and listen at the same time. I don’t feel he served a purpose. He wasn’t live so there’s no personal response.* [#404 in session 1]

*No, because it was not "live." I prefer to speak to my instructors and have them respond.* [#507 in session 1]

*If you consider interaction to be just watching the instructor’s lecture, then I would say that I didn’t mind it. I tend to view interaction as a face to face type process where questions and answers can be verbally put forth if needed.* [#213 in session 2]

*No, I felt that I really didn’t have any. This was more of a one way interaction to me. Some of my questions weren’t answered and so the virtual teacher wasn’t interacting with me.* [#405 in session 2]

*I didn’t feel like there was any interaction. The pop up questions were just automated, not actual interaction.* [#508 in session 2]

Several students mentioned that they did not like the interaction or did not think it was interaction because the answers they got were just repeated parts of the lecture.
No, I like the interaction with instructor in class. You can ask the specific question and get the opinion immediately, without just repeating the lecture. [#504 in session 1]

No, I do not feel like it was interacting. He was just repeating what he had already said earlier. [#312 in session 2]

Some students instead took a more neutral stand on this issue. They felt that the virtual interaction especially QA was helpful, but still was not compatible with the interaction with a live instructor. They said that:

it wasn't bad. it gets straight to the point if that is what you need to be done [#325 in session 1]

I think the interaction the the virtual instructor could be improved. I donot think his answers were as tailored to my questions as they would be in a traditional setting. [#410 in session 1]

'O.K. You can still "talk" to your virtual instructor but it doesn’t make up for the real teacher. [#515 in session 2]

Neutral on this matter also, I don't feel that there is much interaction besides listening to him. Interaction to me is the ability to ask specific questions and have them explained to me in a different manner than the original attempt. [#324 in session 2]

I think that parts of it are good, but there really needs to be more interaction to have an effective learning environment. Meaning, there needs to be more than just a repeat of the lecture. [#423 in session 2]

Finally, although we found that students did not actually interact much with the instructor in the Real Mentor group, most of students still said that they liked the interaction with the instructor, because:

Yes, he was willing to answer questions fully [#608 in session 1]

Yes, because he relates the material to things that are going on right now and shows how you can use the material he teaches in the real world. [#622 in session 1]
The interaction was good, as the instructor was open to any questions, he could answer [#609 in session 2]

There wasn’t much but I do like being able to personally talk with them to answer my questions. [#624 in session 2]

Therefore, it seems that the students’ understanding of interaction in learning was a bit different from what we expected. It was interesting that the potential or capability of interaction was counted as good interaction.

Of course, there were also students in the Real Mentor group who did not have interaction with the instructor and also did not feel the interaction at all. For example,

I didn’t interact with him. I don’t know him, and there were too many people talk with him to get to talk to him. [#607 in session 2]

Question No. 8: Other comments:

Most users did not provide answers to this question. For those who did respond, most of them reaffirmed their positive attitudes toward the system, with comments such as

Overall I liked this form of learning, but I don't think it will work for all courses. [#324 in session 1]

this would be a great addition to students with busy work schedules! [#421 in session 1]

i think many people will like this better than attending school in a classroom. it will save people money and their time.  [#518 in session 1]

I thoroughly enjoyed experiencing this new system of interactive learning. There were many things that I liked and disliked about the system. As I said, a system like this would probably be the most effective with me if I were taking a review type of class. But if I were taking a complex course for the first time, I believe that I would be more successful in a traditional course [#213 in session 2]

I think that this would be a good option for classes, it provides more interaction then the normal online classes but still gives you the freedom to complete it at your leisure. [#403 in session 2]
7. CONCLUSION

To summarize, this dissertation 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 such a “virtual mentor”, which are proposed in the exploratory LVM model.

The key factors proposed in the LVM model include the system interactivity enabled by different information technologies, students’ mental models such as learning style and motivation, and learning phases. The learning outcomes are measured both objectively by test scores and subjectively by students’ perceptions. Taking a multi-methodological research approach, we studied the LVM model by building prototypes, conducting experiments, and observing user experiences iteratively. The relationship among the system effectiveness, learning technologies, students’ mental models, and their learning behaviors has been explored thoroughly in these studies. The results partly support some propositions in the LVM model, but also bring more questions for future research.

7.1. Expectations and Outcomes

We propose that the effectiveness of e-learning is the result of many different factors. Particularly, students’ learning behavior, learning performance, perceived learning effectiveness, perceived system usability, self-report interactivity, and preference of e-learning are all positively related to the level of system interactivity (assuming real mentor has the highest level of interactivity), the learning phase, the student’s intrinsic
motivation, capability of cognitive strategy, and active learning style. We also propose that these learning outcome measures may be negatively related to the student’s extrinsic motivation, and sensitive or sequential learning style. Different information technologies enable different levels of system interactivity. We propose that the virtual interaction in form of video-based question answering provides higher interactivity than the hypermedia type of interaction (navigation through hyperlinks in a lecture outline). Their combination, as well as the combination with automatic pop-up quizzes, provides even higher level of interactivity.

Results from experiments and observations showed that, first, the virtual interaction technology, learning phases, and students’ learning styles did impact the students’ learning behavior to some extent as we expected. All groups had more interactive learning activities in the advanced learning session. Particularly students who were only provided the QA function had significantly more interactive activities than those who were only provided the outline function in the advanced learning session; however, more system interactivities (the combination of QA and outline, as well as the combination with pop-up quizzes) did not significantly increase further the actual learning activities in this session. This might be due to the students’ activities of answering the pop-up questions in the first session not being recorded in the system logs, the pop-up questions not being included in the second session, and therefore the quiz activities not being counted. We argued that it was the QA function that triggered more interactive learning activities in the advanced learning session. However, more studies need to be done to confirm this argument.
Surprisingly, we found that there were very few interactions between the students and the instructor in the classroom. A follow up interview revealed that many students were reluctant to ask questions in the classroom because of concerns such as being viewed as stupid, disturbing the whole class, or waiting to get answers (in the advanced learning session). We found that virtual interactions could actually remove these psychological barriers and enable more direct question and answering processes in learning.

However, more surprisingly, the students’ perceived (self-reported) interactivity is very different from their actual interactive learning activities. First of all, the perceived interactivity was not significantly different between sessions. Therefore the students’ perception about interaction did not change along with their actual behavior. Secondly, the perceived interactivity was significantly higher in the real mentor group than in the other five treatment groups both in the advanced learning session and in the overall learning process. Therefore, 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 interacting. The qualitative results confirmed this finding. On one side, the students’ understanding 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 non-verbal communication features such as eye contact or gestures, as well as the feeling of social presence, might contribute to this circumstance. For example in a real classroom, even if a student does not interact with the instructor but
others do, he or she can still view the instructor as interactive. But with LVM, the
interactivity all has to relate to the student him or herself. If he or she does not personally
interact, then the system is faulted; whereas in a real classroom, the onus shifts to the
instructor, who can satisfy the interactivity requirement by interacting with any one of the
students. 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 pre-recorded lecture. Therefore, most students quickly realized that virtual
mentor could not rephrase the answer based on their individual need; neither 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. We think that this limitation
might be relieved by extending the answers with contents extracted from the Web.
However, the social side of the interaction with a live instructor will be difficult to be
realized by virtual interaction with current technology. Asynchronous collaboration may
still be needed to complement the virtual interaction.

Although the students in the virtual mentor groups did not feel as much interaction as
those in the real-mentor group, different level of system interactivity did impact the
perceived interactivity, in a way consistent with the actual interactive learning activities.
Students who were only provided the QA function had significantly higher perceived
interactivities than those who were only provided the outline function, both in session 2
and in the overall learning process. The qualitative results confirmed this finding again,
in that students’ satisfaction towards outline decreased from session 1 to session 2, while
their satisfaction towards 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. Overall, the QA type of virtual interaction did enable more actual interaction and perceived interactivity than the traditional hyperlink type of interaction when students need more interactions in the advanced learning session.

In terms of the relationship of students’ characteristics and their interaction in learning, we confirmed that the more active type of learners (higher score on ACT) would have more actual interactive activities in the advanced, task-oriented learning process. However, those more active types of learners had less perceived interactivity. One possible explanation to these unexpected results is that active learners had more expectations on interactions than what the virtual mentor could provide. This discrepancy generated their dissatisfaction towards the virtual mentor.

The results about the impact of virtual interaction on the learning outcomes are also very interesting. A more powerful ANCOVA test revealed that there was no significant difference among system groups found in either the second session or the overall learning process. But the real mentor group, which was assumed to have the highest learning performance gain in both session 2 and overall, actually had the lowest learning performance gain. This is, however, consistent with our observations that the students had the least actual interactive learning activities in the real mentor group in both sessions.
Again, similar with the findings about interactivity, the students’ perception is the opposite of the reality. Overall, the perceived learning effectiveness was significantly higher in the real mentor group than in the other five treatment groups, although we again failed to find significant difference among the five treatment groups. However, the perceived learning effectiveness and the actual learning performance was not significantly correlated in each session.

Since we did not find any significant correlation among the actual interactive learning activities, the learning performance, perceived learning effectiveness, self-report interactivity, and learning system preference, we conclude that although the virtual interaction technology can directly impact the students’ learning behavior, the change in interaction behavior may not directly affect the learning performance and students’ perceived learning effectiveness. Particularly, although students viewed QA as better virtual interaction than outline, their more interaction behavior enabled by QA did not result in significantly better learning results. Therefore, interactivity did not play a key role in determining the effectiveness of learning, although we did find that e-learning systems provided more effective learning in the advance learning session. Looking at the results from qualitative observations, we realize that the self-paced control and the 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 determines different level of convenience of reviewing specific content, but as long as this convenience exist, students can learn better with virtual
mentor than with real mentor. The level does not matter so much to the learning effectiveness. The benefits are more from the satisfaction perspective. For example, when QA functions were available, students did include QA as one of the reasons why they liked the system; while only a few students included the outline or the pop-up quiz as one of the reasons. In addition, there were twice or three times more students in the second session listed QA as one of the reasons they liked the system than in the first session. This again confirmed that interaction especially QA is more important for the advanced learning phase.

In addition, the discrepancy between reality and students’ perceptions suggests that most students hold a stereotype that learning with a real mentor face to face is always the best, which is not necessarily true in the advanced learning phase. Both the quantitative and qualitative results about students’ preference on learning systems again confirmed this finding. Most students said they preferred learning in traditional classroom settings over learning with a virtual mentor system, and this preference did not change across session or group. Again, the social involvement feeling might be the major reason for this preference, but more experience with the virtual mentor may also change this stereotype. A longitudinal study is needed in the future.

In terms of the impact of students’ characteristics, we found that students who were less extrinsic goal driven and students who were more intuitive type of learner (lower score on SNS) would have more overall learning gain when learning with virtual mentors. However, although more active type of learners did have more interactive activities in the advanced session, they had less perceived learning effectiveness and preference on
learning systems. This again might due to the discrepancy between the active learners’ high expectations on interactions and the current limitation of the virtual interaction.

About perceived system usability, except the group with the combination of QA and outline had increased perceived system usability, all the other groups had decreased perceived system usability from the introductory learning session to the advanced learning session. This implies that students thought the combination of QA and outline was the easiest way in finding the specific information.

Finally, we found two factors that were possibly missing in the current LVM model from the qualitative observations. First, the instructor’s teaching style might greatly influence student interactions in learning and their perception of the effectiveness of e-learning. Some students did not like the system or the lecture just because they thought the instructor was annoying and the lecture was boring. However, it seemed that the instructor could unintentionally change his behavior when being videotaped. Therefore, it is would be interesting to study how different instructors and their different behaviors might affect students’ learning outcomes in e-learning. Secondly, many students pointed out that the LVM system might only benefit those lecture-based courses teaching about facts. For more hands-on courses or courses need a lot of discussions, the virtual interaction would not be sufficient. Therefore, when studying the effectiveness of e-learning, the type of courses may also be an important influencing factor.

To summarize, 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. However, its impact on the learning effectiveness is
limited. The originally proposed LVM model need to be revised by adding more factors such as “pace control”, “convenience for access information”, course type, lecturer style, and probably a factor about students’ preferred level of social involvement.

7.2. Implications

The research findings reported in this dissertation have both research and practical implications. From the research perspective, this dissertation clearly demonstrates that the effectiveness of e-learning is influenced by many factors besides the information technology. Through both quantitative and qualitative studies, the key factors are identified and some relationships among these factors are discovered. Although the original proposed conceptual model LVM is not completely supported, this dissertation provides evidences on how to revise it. We hope this revised model will become a theoretical foundation for future development and deployment of effective e-learning.

In addition, this dissertation studies the effectiveness of e-learning with a focus on the impact of interaction, particularly a new type of interaction defined as virtual interaction. The virtual interaction is still learner-content interaction but it simulates the learner-instructor interaction. We believe this virtual interaction will bring students more convenience in accessing information than traditional hypermedia type of learner-content interaction. It should also improve students’ satisfaction with e-learning because it gives students feelings of talking with a live instructor. The virtual interaction does not rely on the availability of human instructors and therefore will be more cost effective, reducing instructors’ work load and allowing more flexibility in learning. Although virtual interaction is found not to be able to directly impact the learning performance, it is
somewhat related to students’ satisfaction toward 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 dissertation provide suggestions for e-learning practitioners. Since the effectiveness of e-learning is influenced by many different factors, when deploying e-learning in reality (e.g. providing e-learning courses in college), all these factors need to be considered besides simple technology features. For example, students need to test themselves to know if their learning style and motivation will be suitable for e-learning. The technology features of e-learning need to be adjusted based on different learning phases and course types in order to achieve the best learning outcomes. Specifically, for lecture-type of courses or introductory courses, using simple hypermedia lectures is fine and can reduce costs. In advanced learning phases that reviewing and reinforcing previous content is important, the virtual interaction can be added to help students obtain answers with reduced waiting time, and at the same time reduce the instructor’s work load. However, based on the observations in this dissertation study, virtual interaction may not be sufficient for more advanced, discussion type of class, and ALN or face-to-face classroom discussion will still be needed. The appropriate combination and complementary of virtual and real mentor will be the best way to learn.

7.3. Limitations and Future Research

The exploratory nature of this dissertation study determines the existence of certain limitations. First of all, the internal validity of the LVM model was challenged when we found that course type and teaching style might also be key factors affecting the
effectiveness of the LVM system. However, since we had only one instructor teaching one lecture-type of course in the experiment, these factors were controlled so that the results about the other proposed factors were still valid. In the future, the LVM model will be revised to incorporate these new factors and new experiments will be conducted to test them. Particularly, more different types of lectures taught by different instructors (or the same instructor in different styles) need to be recorded and added into the LVM system.

Another limitation of this dissertation study is about the use of experimentation and observations together. Although experimentation allows us control the constructs more easily, the limited access (fixed time and fixed location) is very different from the real scenario of how a student may use a Web-based learning system such as the LVM system. Therefore, although it is more convenient to get the observations in and after the experimentation, and the observations can help interpret the experimentation, such observations may be biased by the limited access. To observe how students really use this system in a natural setting, a field study or case study is needed. Therefore in the future, besides continuously conducting controlled experiments to validate the revised LVM model, we will conduct a separate field study to observe students’ behavior and perceptions in a more realistic setting.

The current study is also limited by the amount of contents available in the system and the length of the experiment. In the current experiment, we only had a one-hour long lecture available in the system, and the experiment had only two sessions, each of which was about 2 hours long. In such a short time, students might not be able to fully
experience the advantage of the system. Particularly, for a lecture only about one hour long, students could find information by simply watching through the video. In addition, many students would find their questions have to be limited to the specific lecture. If there were a large collection of lectures, manual searching would be impossible, the QA type of virtual interaction would be critical, and answers for the student’s question would be extended. Therefore, students’ behavior and perceptions about the LVM system may also change when there are a large collection of lectures, for example, lectures for one course in a whole semester. Therefore, a longitudinal field study using the LVM for a course across a whole semester is ideal for exploring how the effectiveness of LVM changes over time.

Finally, the current study is limited to the context of college education. Although we think that the virtual interaction will be a good way to provide on-the-job training to professionals, we have not tested it yet in an industrial setting. In the future, we will try to apply the LVM model and system to professional trainings and compare the findings to our current findings from the educational setting. We hope that our study will help bringing effective e-learning to both schools and professional organizations.
## APPENDIX A. CODEBOOK FOR OPEN-ENDED RESPONSES ANALYSIS

<table>
<thead>
<tr>
<th>Record Number</th>
<th>Code</th>
<th>Brief Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3PANEL</td>
<td>like the combination of three panels</td>
</tr>
<tr>
<td>2</td>
<td>TLIMIT</td>
<td>do not like time limit in finishing the lecture</td>
</tr>
<tr>
<td>3</td>
<td>NONE</td>
<td>none or n/a or nothing or maybe…</td>
</tr>
<tr>
<td>4</td>
<td>NOCLASS</td>
<td>when having no access to traditional classes</td>
</tr>
<tr>
<td>5</td>
<td>TRADIT</td>
<td>prefer traditional class</td>
</tr>
<tr>
<td>6</td>
<td>PERSON</td>
<td>like the personal feeling with teachers and students</td>
</tr>
<tr>
<td>7</td>
<td>QA</td>
<td>prefer asking questions</td>
</tr>
<tr>
<td>8</td>
<td>EXACT</td>
<td>can get exact answer</td>
</tr>
<tr>
<td>9</td>
<td>LIKEINT</td>
<td>like the interaction</td>
</tr>
<tr>
<td>10</td>
<td>REAL</td>
<td>the virtual instructor looks like a real professor</td>
</tr>
<tr>
<td>11</td>
<td>GREAT</td>
<td>like the system overall</td>
</tr>
<tr>
<td>12</td>
<td>OWNPACE</td>
<td>like self-paced learning</td>
</tr>
<tr>
<td>13</td>
<td>REVIEW</td>
<td>like being able to review</td>
</tr>
<tr>
<td>14</td>
<td>ANNOYIN</td>
<td>the instructor is annoying</td>
</tr>
<tr>
<td>15</td>
<td>CONFUSE</td>
<td>do not understand the question</td>
</tr>
<tr>
<td>16</td>
<td>ENTRY</td>
<td>use this system for entry level class</td>
</tr>
<tr>
<td>17</td>
<td>SYSTEM</td>
<td>prefer this system</td>
</tr>
<tr>
<td>18</td>
<td>BOTH</td>
<td>no preference on system or traditional class</td>
</tr>
<tr>
<td>19</td>
<td>OUTLINE</td>
<td>prefer to use outline</td>
</tr>
<tr>
<td>20</td>
<td>NOINT</td>
<td>do not feel interaction</td>
</tr>
<tr>
<td>21</td>
<td>EASYUSE</td>
<td>it is easy to use</td>
</tr>
<tr>
<td>22</td>
<td>NOFOCUS</td>
<td>it is easy for student to lose attention</td>
</tr>
<tr>
<td>23</td>
<td>FORSOME</td>
<td>this system will only work for some students</td>
</tr>
<tr>
<td>24</td>
<td>SHORT</td>
<td>use it for short classes</td>
</tr>
<tr>
<td>25</td>
<td>INTEGRA</td>
<td>like the integrated system</td>
</tr>
<tr>
<td>26</td>
<td>LOTINFO</td>
<td>the system offers a lot of information</td>
</tr>
<tr>
<td>27</td>
<td>TEXT</td>
<td>improve text</td>
</tr>
<tr>
<td>29</td>
<td>ONLINE</td>
<td>use it for online classes</td>
</tr>
<tr>
<td>30</td>
<td>ANY</td>
<td>use it in any situation</td>
</tr>
<tr>
<td>31</td>
<td>REPINT</td>
<td>do not like the virtual interaction because it just repeats the materials</td>
</tr>
<tr>
<td>32</td>
<td>CONSTRA</td>
<td>use the system in constraint situation</td>
</tr>
<tr>
<td>33</td>
<td>QUICK</td>
<td>can get information quick</td>
</tr>
<tr>
<td>34</td>
<td>OK</td>
<td>it is just ok, not like, but not dislike</td>
</tr>
<tr>
<td>35</td>
<td>OVERLOA</td>
<td>too much on screen, cognitive overload</td>
</tr>
<tr>
<td>36</td>
<td>NOCHOIC</td>
<td>when there is no other option</td>
</tr>
<tr>
<td>37</td>
<td>NEWWIN</td>
<td>do not like the new pop-up window when using the ask question function</td>
</tr>
<tr>
<td>38</td>
<td>FACT</td>
<td>use it for factual information</td>
</tr>
<tr>
<td>39</td>
<td>BETTERA</td>
<td>the human instructor can provide better answer</td>
</tr>
</tbody>
</table>
QALIMIT limitation with QA function
LECTURE use it in pure lecture
BORE it is boring
BETTERO better structured outline
EYETIR eyes tired on screen
NOTES when I need to take notes
COMBINE have the virtual lecture combined with traditional
AV sound or video
RELATE can relate the information back to everything else
APPEAR concise appearance, take less screen space
SPEC specific classes (e.g. MIS courses, math courses)
RETHEN better retention
INSTR provide information about instructor
NODISTU do not interrupt the whole class with questions
WANTIT want the system to be implemented soon
NOPOP no pop up quizzes
READ prefer reading text or only reading
EXPAND provide more information, e.g. hyper links
EMAIL use email for interaction
RESTART technical problem with QA, have to restart
HOME use the system at home
NOCARE do not care
POPUP use popup questions
ASSIGN like the assignments
PREVIEW use this lecture for a preview before real lecture
PC have to deal with a computer
MOREINT add more interaction
COMMENT provide place for students to comment and feedback
LIVE make the lecture live with microphones
USEDTO get used to the traditional way of learning
ONJOB work related
DISCUSS provide function for students discussion
SYNC synchronization of the slides
HANDSON no hands on examples
EXAMPLE instructor give examples
CONTENT like the content of the lecture
REPEAT content was taught before
TEACHER the teacher is knowledgeable
SLIDES like the slides
ORGANIZ the lecture is organized
FEEDBK provide feedback to the quiz questions
FWRW fast forward / rewind buttons
INTERES something to add interest in the lecture
REFERENCES


Gallagher, S. "Online Distance Education Market Update: A Nascent Market Begins to Mature," Eduventures.


