* * *

Developing an Interaction-Centered Evaluation Tool for Distance Education

MASTER OF SCIENCE THESIS

* * *

John Perrine, MD

SCHOOL OF MEDICINE
DEPARTMENT OF MEDICAL INFORMATICS AND CLINICAL EPIDEMIOLOGY

OREGON HEALTH & SCIENCE UNIVERSITY

November 2003

* * *

SCHOOL OF MEDICINE

DEPARTMENT OF MEDICAL INFORMATICS & CLINICAL EPIDEMIOLOGY

* * *

CERTIFICATE OF APPROVAL

* * *

THIS IS TO CERTIFY THAT THE M.S. THESIS OF

JOHN PERRINE, MD

HAS BEEN APPROVED

[redacted]

WILLIAM HERSH, MD

[redacted]

KATHERINE O'MEARA, RN, PHD

[redacted]

DALE KRAEMER, PHD

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to acknowledge the generous assistance of my thesis advisor, Judith Logan, MD, and the guidance of my thesis committee members: William Hersh, MD, Dale Kraemer, PhD, and Katherine O'Meara, RN, PhD. Thank you.

I would also like to acknowledge the instructors who served as expert reviewers of the evaluation tool, and provided valuable insights about distance education: Joan Ash, PhD, Paul Gorman, MD, Zoë Starve, PhD and Christopher Dubay, PhD.

Many thanks for the moral support of Pat Tidmarsh JD, MS, Dan Baker, MD, Tadaaki Hiruki, MD, Matthew Kavanagh, MS, Pramod Jacob, MS, and Andrea Ilg.

Thanks to my mom and dad.

ABSTRACT

OBJECTIVE: Web-based distance instruction is growing in usage and importance. Coming with this growth are changes in instructional methods and technologies. There is a need for tools to evaluate these new courses, instructional methods and technologies. One possible strategy is evaluation based on student interaction. The objective of this project was to produce an interaction-centered evaluation tool for distance education.

METHODS: Question items were developed for an interaction-centered evaluation tool based on a review of distance education literature and interviews with Oregon Health & Science University distance education faculty. Fifty-five students in seven medical informatics courses participated in a measurement study to determine the instrument's validity and reliability. Questionnaire items were included or rejected based on corrected item-total correlation and correlation with questions thought to measure the instruments underling constructs. This selection process also served to demonstrate the instrument's criterion-related validity. Internal-consistency reliability and test-retest reliability were also determined for the instrument.

RESULTS: The resulting 36-question interaction-centered instrument was designed for summative or formative evaluation of distance courses and course components. This instrument was used to analyze and make recommendations about two of the informatics courses used in the measurement study. Analysis scores and student comments suggested

modifications in course design and instruction methods to improve student interactions within courses.

CONCLUSION: The Web-based evaluation tool developed in this project measures student perception of quality of course interactions with demonstrated validity and reliability. It is easy to use and delivers intuitively meaningful information about distance courses.

INTRODUCTION

Web-based instruction is a new and rapidly growing form of distance education. Sixty-two percent of American colleges and universities now offer online courses, compared with 57% in 2001 and 46% in 1999 (Campus Computing Survey, 2002). Universities across the world offer Web-based courses to millions of students. The University of Phoenix (2003) currently claims 57,000 online students in addition to its 84,000 classroom-based students, while the UK's Open University (2003) claims over 75,000 online students.

The move to Web-based education is also occurring in the health sciences. In 2002, Oregon Health & Science University began a Web-based Masters program in medical informatics (OHSU, 2003). In 2003, Stanford University began a Web-based Masters program in bioinformatics (Stanford University, 2003). Brown University in Providence, R.I. is part of an international consortium developing an online curriculum for the first two years of medical school. This consortium plans to start piloting the International Virtual Medical School this fall at the University of Dundee in the UK (Forelle, 2003).

A number of factors are supporting this revolution. The declining costs of computer and telecommunication technologies are lowering barriers to student participation. Access to the Internet is allowing easier and faster communication between students, instructors and course content. Students also have an expanding number of courses, programs and institutions to choose from.

What's more, the quality of instruction may be improving as distance programs adopt new technologies and new instructional methods (DiSessna, 2000). Some educators, however, feel that distance education is inadequate compared to classroom education: Eighty-two percent of accounting professors responding to a 1999 survey thought that distance-education suffered from a lack of adequate student-student and student-instructor interaction (Saunders and Weible, 1999).

Given that this technology is very young, the quality of online education has not been well-studied. Is the quality of online education improving? Is it even adequate? As technologies and instructional methods change, educators need tools that evaluate course performance and tools that aid in course design and development. Examples of questions illustrating summative and formative evaluation:

- *Summative evaluation*
  - How does the course compare to accepted standards?
  - How satisfied are students taking the course relative to students taking other courses at the institution?
  - What is the course's quality relative to other courses at other institutions?
  - How does course quality affect outcomes (e.g. grades, student retention)?
- *Formative evaluation*
  - How can technology be utilized in the course?
  - Which services, processes and/or technologies, do educators need to institute?
  - How can costs be lowered?

- How can access to courses be improved?

These questions may look familiar because they are similar to questions medical informaticians ask about the quality of healthcare. With an outcomes-based approach to quality improvement, healthcare should improve, cost less and provide greater access. These criteria should also be true for education. If the outcomes of education are studied, and quality improvement techniques implemented, then all aspects of students' education should benefit. Improving evaluation tools for distance education is part of this process, and was the intent of this thesis project.

Distance education students

To a large degree, students participate in distance education because of its convenience and accessibility. Many students participate in distance education because they live too far away to attend classroom-based courses. In addition, students living close to campus may take distance classes for the flexibility of anytime, anywhere access (Northrup, 2002). Distance students also tend to prefer independent learning and/or a more conceptual style of instruction (Diaz, 1997).

But are distance classes as good as those in the traditional campus setting? Russell's *"No Significant Difference"* collection of 355 studies examines learning outcomes, grade comparisons and satisfaction surveys for distance and computer-mediated education (Russell, 1999). This collection of studies provides evidence that technology-mediated and distance courses perform as well as their on-campus counterparts. Some examples:

- A 1999 study of graduate engineering students found no significant difference in final exam scores between students in traditional lectures and students viewing lectures online (Kortemeyer and Bauer, 1999).
- A 2000 study of graduate students found an equal distribution exam scores and no significant difference in satisfaction between online and classroom students. Independent judges in the study also found no significant difference in the quality of major course projects between the two formats (Johnson, *et al.*, 2000).
- A 2001 study of graduate accounting students found no significant difference in student performance between online and classroom formats. Online students, however, were less satisfied with instructor availability than classroom students (Gagne and Shepherd, 2001).

Brown and Duguid (2000) argue that while learning outcomes may be similar, online education may neglect the social aspects of learning. On campus, students interact with classmates and instructors in the classroom and in social settings. Opportunities for interaction on campus are many and varied, from comparing notes and participating in study groups to eating pizza and going to football games. Palloff and Pratt (1999) believe that student interaction is the heart, and pedagogically most valuable, aspect of the educational experience. In distance education these interactions can be challenging to accomplish, and may not be happening at all. It is important, therefore, to monitor and facilitate communication in distance settings.

Constructivism in distance education

Constructivism has its theoretical roots in epistemology, the philosophy of learning or knowing. It is the theory that knowledge is constructed through the process of exploration: Individuals construct their own knowledge, opinions and understandings based on exploration and interaction the environment. English philosopher Locke conceived that the mind was imbued with simple, innate ideas around which knowledge is formulated by interaction with the world. German philosopher Kant also believed that the mind started with built-in categories, or ideals, which provided spatial, temporal and causal frameworks on which new and more complex ideas or frameworks could be built. Knowledge is then acquired by filtering perceptions through these mental frameworks (Oxford, 1997).

In constructivist learning theory, students acquire knowledge through active inquiry, problem solving and collaboration. Constructivist theory challenges the more traditional *transmission* model of learning, in which the instructor is a repository of knowledge that is passed on to students in the form of lectures and dialog. In constructivist theory, the instructor a guide who encourages students to explore, collaborate, and form their own ideas and conclusions (Abdal-Haqq, 1999). While constructivist educators consider lecture and memorization to be important methods in education, they would also argue that passively acquired knowledge is not integrated as well into durable knowledge structures as is knowledge (Richardson, 1997).

In education, there are two schools of constructivist theory. *Individual (Piagetian or psychological) constructivists* regard learning as a knowledge-construction process performed by individual students. *Social constructivists* view knowledge creation as a process that takes place in a community, or within an individual interacting with community members.

In *The Construction of Reality in the Child*, developmental psychologist Piaget (1954) introduced the theory of (what is now referred to as) individual constructivism. Piaget saw the child as an independent explorer with natural curiosity and a set of simple conceptual frameworks, or *schemata,* around which new knowledge could be built: As the child interacts with his environment, these schemata become more developed and complex. These are the processes of *discovery* or *acquisition.* Discovery of new information leads to *cognitive disequilibria* that allow for the acquisition of new knowledge (Oxford, 1997).

Russian psychologist Vygotsky is credited with the development the theory of social constructivism, which places learning within a social and cultural context: Knowledge construction is a collaborative process in which new meanings and understandings are developed through dialog, negotiation and consensus. Individual learning, therefore, takes place through interactions with community members, and cannot be isolated from social and cultural influences. In the classroom, students adopt social and cultural norms through group activities and the examples of teachers and classmates. This interaction is a two-way street: As the student is learning from the community, the community is learning from the student.

Vygotsky, like Piaget, believed that learning took place around mental frameworks. He stressed that these *scaffolds* could be introduced and/or reinforced by an instructor and removed when they were no longer needed. Since students learn at different rates and approach learning with different ideas, expectations and needs, Vygotsky stressed the importance of differing avenues of knowledge construction. He termed these different avenues *zones of proximal development*: places where students could get the needed guidance and support from teachers and other community members (Oxford, 1997).

Several instructional methods are suggested by constructivist theories: learning should be active, should contain interaction with community members, and should be relevant and learner-centered.

- **Learning should be interactive:** Learning should include interaction with community members. Interaction with instructors and other students allows construction of shared knowledge and provides motivation for learning. What's more, interactions with persons holding differing knowledge bases, resources and viewpoints can provide learners with different frameworks for understanding (provide different zones of proximal development). Collaboration with peers of differing capabilities and skill sets can create environments that promote critical thinking, negotiation and social skills sets. Collaboration also creates an emotional/intellectual climate that promotes the formation of learning communities. Tu and Corry (2002) describe these communities as groups that learn by defining problems, negotiating solutions and collaboratively taking action. As these communities interact, they gain knowledge and skill. Community membership encourages analysis of ideas, multiple viewpoints, multiplicity of feedback and mutual guidance.

- **Students should actively participate in leaning:** American education theorist Dewey advocated problem-based, active learning. Dewey felt that acquisition of knowledge was accomplished through *discovery* – a process that is facilitated by encountering and solving real-life problems. In order for problems to be relevant and motivating to students, they should be based on the students' interests and real-life experiences. In terms of individual constructivist theory, students should be confronted with dilemmas that challenge their schemata (Abdal-Haqq, 1999). Through confrontation, students' ideas, opinions and beliefs can be buttressed, modified or deconstructed.

- **Course materials and learning activities should be relevant:** Students' interests and experiences should drive learning. Course materials and learning activities perceived as relevant and interesting provide motivation for learning and allow for integration of knowledge into existing knowledge structures. Adult learners want to acquire knowledge and learn skills that meet their life experiences and work needs: Instruction, therefore, should be rooted in meaningful problems, events and issues. Relevant activities such case studies, simulations and internships should be included in course designs (Huang, 2002).

- **Learning should be student-centered:** As opposed to the instructor-centered transmission model, the constructivist model of education has the learner as its focus. Educators should be facilitators or guides, not just transmitters of knowledge. To the degree possible, courses of study should be individualized for the needs and abilities of each student.

Constructivist methods have been adopted by many educators and have become common in today's classroom. Distance educators and designers, however, face challenges in adopting constructivist principles. Before the availability of Web-based and computer-mediated methods, the use in distance education of constructivist methods such as interaction, collaboration and active learning were difficult. Correspondence and broadcast modes of communication tended to reinforce the transmission model of education (Rumble, 2001). Little more was possible with those limited modes of communication. The Internet has made it possible for students to enter into dialogs, collaborate, and access learning materials from multiple sources. As broadband Web access becomes more common, constructivist methods should increase in distance education, and because adult distance learners tend to take control of their own learning, constructivist methods may also be a better fit than transmission-model methods.

Moore's interaction model

Constructivist theory suggests that students should be active participants in learning, and that student interaction should be a principle component of education. Moore (1989) proposed a model of distance education based on three categories of interactions: *student-instructor, student-student* and *student-learning materials.*

Moore suggested instructors should interact with students to stimulate the student's interest in learning topics and motivate the student to learn. This should be accomplished through organization of learning applications and practical skills, as well as feedback, support and encouragement for each student. The instructor should also interaction with students to elaborate on subject matter, clarify individual misunderstandings, and suggest supplementary study. He proposed that student-student interaction was also key to stimulation of interest and motivation, especially in younger learners.

Moore considered student-learning material interaction to be fundamental: "Without it there cannot be education, since it is the process of intellectually interacting with content that results in changes in the learner's understanding, the learner's perspective, or the cognitive structures of the learner's mind … when learners 'talk to themselves' about the information and ideas they encounter in a text, television program, lecture, or elsewhere."

These student-learning material interactions should be designed to motivate learners, facilitate application of learning principles, and provide affective support to students.

Hirumi (2002) extended this model to include student-instruction interaction: interaction with instructional methods such as examinations (Figure 1).

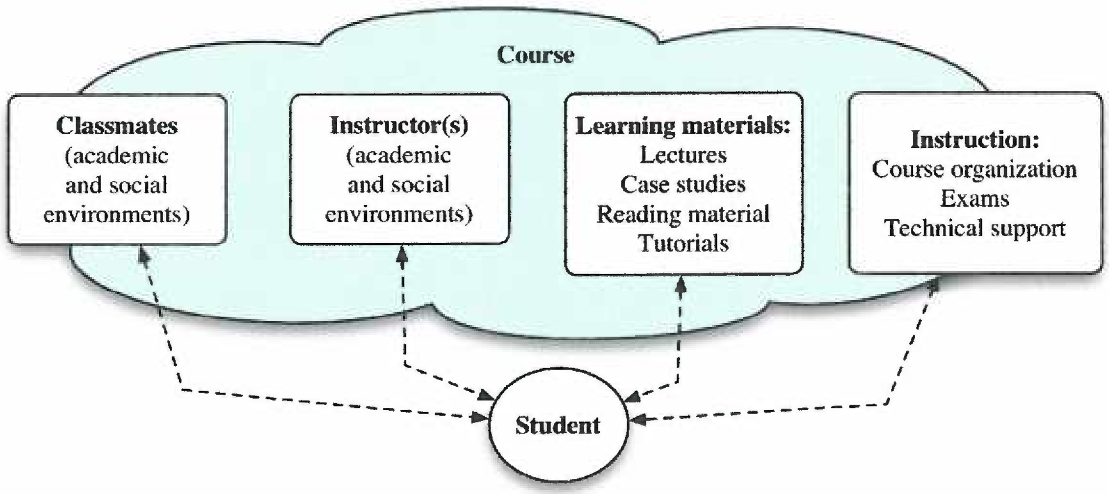


Figure 1. The Moore-Hirumi interaction framework – distance courses are made of student interactions with classmates, instructors, learning materials, and instructional methods.

Interaction in distance education: research findings

Research in distance education supports the instructional methods suggested by constructivist theory – specifically the benefits of student interaction suggested my Moore and Hirumi.

Student-classmate interactions

Collaborative learning is beneficial in face-to-face education (Johnson, 1981). Recent research has demonstrated its success in distance settings. In a triangulated (qualitative and quantitative) study at New Jersey Institute of Technology, Hiltz (2002) found that collaborating distance students obtained outcomes (grades, perception of quality, satisfaction) as good or better than their on-campus counterparts while distance students who did not collaborate obtained outcomes poorer than on-campus students.

In a qualitative study on law students, Parry and Dunn (2000) found that collaborative distance learning fosters a more thorough understanding of concepts and ideas. Investigators also found that assessment of group tasks encouraged students to employ higher-level thinking and to learn skills from each other. Parry and Dunn note that although some law students found group projects to be onerous, collaboration was a good motivator. Students also reported that collaboration was an enjoyable experience and essential to the quality of their learning.

Student-instructor interactions

In a qualitative study at Arizona State University, McIsaac (1999) found that quality and timeliness of instructor feedback impact learning outcomes. Because students in distance settings lack many of the feedback and social-presence cues of the classroom, such as facial expressions and gestures, timely feedback from instructors is important.

A University of West Florida questionnaire survey of graduate students examined preferences for distance interaction (Northrop, 2002). Investigators found that instructor availability and timeliness of response were important to student satisfaction. Investigators also found that too many interactive assignments caused frustration. Berge

(1999) concurs that student-instructor interaction is important to student satisfaction and motivation, but too much interaction may lead to frustration, boredom and overload.

Student-learning material interactions

Levin *et al.* (2002) hypothesized that multiple representations of knowledge benefit students by allowing them to examine problems and issues from different points of view, and allowing students to select approaches, media and methods that work best for them. Levin's qualitative study demonstrated that students prefer multiple modes of media, instructional formats (e.g. lectures, textbook readings, student presentations, simulations) and learning activities to choose from. Northrop (2002) also found that students prefer a variety of instructional techniques, and that students are most comfortable with the feel of a traditional class -- lecture and readings followed by discussion.

Laurillard and Taylor (1994) in a study conducted at Open University in the UK found that high-quality sound and visuals are important factors in student motivation and attention. Their study found that video is beneficial in its ability to convey vicarious experience and conceptual understanding. They also found that student interactivity is important in maintaining productivity.

Clark (1992) argued that learning objectives are largely independent of the type of media. He concluded that achievement gains result from the use of media and methods that are appropriate to learning objectives and student aptitudes: Examples, analogies, models, interactive simulations and feedback all enhance learning, regardless of the media of implementation. He reasoned that since the vehicle of presentation has little impact on learning, media should be selected to maximize student access and cost-benefit. He suggested that routine information could be presented using less expensive media, and that expensive media should be reserved for more challenging information and concepts. A randomized controlled trial comparing the value of video to audio lectures (Berner and Adams, 2003) supports Clark's argument. This study of graduate-level distance students found that video lectures demonstrated no advantage over audio lectures in either satisfaction or learning outcomes.

There is agreement between the theory and the research of distance education that interaction in learning is important, and the quality of these interactions affect learning outcomes. Table 1 summarizes findings and recommendation

Table 1. *Summary of instructional criteria suggested by distance education theory and research*

Student-classmate interaction:	**Student-instructor interaction:**	**Student-learning material interaction:**	**Student-instruction interaction:**
• Collaboration • Social connections • Academic networking • Formation of learning communities	• Student centered • Individualized • Feedback: Reinforce concepts Answer question Clarify misunderstandings • Motivation and interest • Academic networking	• Active learning • Student centered • Interesting • Relevant • Multiple representations of knowledge and modes of instruction • Stimulate critical/creative thinking	• Promote learning • Facilitate goal achievement

Consortia Recommendations

Other support for the importance of student interaction in distance education comes from the American Distance Education Consortium's *Guiding Principles for Distance Teaching and Learning* (ADEC, 2002), and the American Association of Higher Education's *Principles of Good Practice in Undergraduate Education* (Chickering and Ehrmann, 1996). These groups suggest educational principles deemed necessary for quality distance education (Table 2).

Table 2. *Distance education consortia recommendations*

American Distance Education Consortium *Guiding Principles for Distance Teaching and Learning*	**American Association of Higher Education** *Principles of Good Practice in Undergraduate Education*
• Student-classmate interaction • Support development of communities • Active learning • Problem-based and knowledge-based learning • Appropriate use of a variety of media • Clear purpose • Tightly-focused outcomes and objective • Contribution to a social mission of education in a democratic society	• Interaction between the student and instructor (tutor, TA, mentor or expert) • Student-classmate interaction • Active learning • Adequate time on task for learning activities • Substantive, rapid feedback • High expectations of instructor • Respect for different talents, abilities of students

Interaction-centered course evaluation

The interaction framework based on the work of Moore and Hirumi suggests an evaluation strategy for distance education (Figure 2). From the distance student's perspective, course interactions *define* the course. Therefore, the quality of a course depends on the quality of course interactions.

A tool that measures the quality of interactions would have the potential to improve distance courses by providing summative information about quality and formative information about how interactions may be improved. The purpose of this project and study was to develop and validate an interaction-centered evaluation tool for distance courses. The tool was designed to measure student perception of interaction in distance education in such a way as to be:

- Short and user friendly
- Intuitively meaningful
- Valid and reliable
- Easily used across a variety of applications
- Easily modified over time

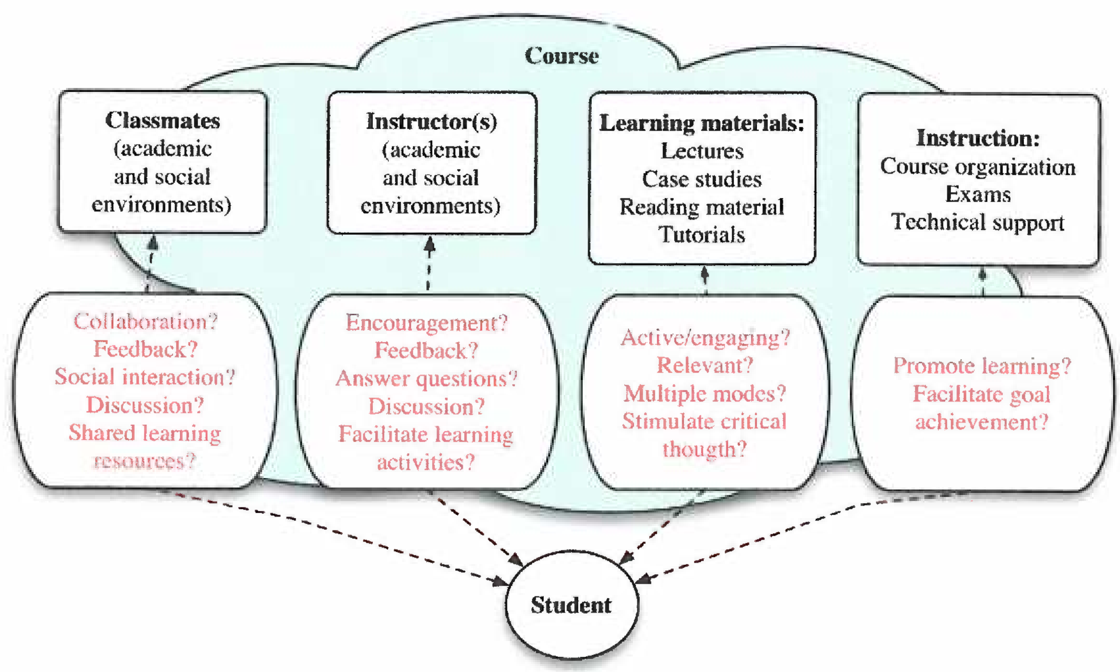


Figure 2. An evaluation strategy based on the Moore-Hirumi interaction framework – categories consist of different interactions (red text) that may be evaluated. The quality of the individual interactions determines the quality of the course.

Evaluation of distance education

Some studies of distance education have used evaluation methods designed for on-campus instruction. These tools have not been developed for measuring distance education. The Purdue Rating Scale (Derry, 1974), for example, has been used in distance education studies (Anderson, 2002), though its validity and reliability has not been demonstrated in distance settings. Similarly, the tool used to evaluate distance courses at OHSU in the Department of Medical Informatics & Clinical Epidemiology was developed for classroom-based courses, and adapted for distance courses. This summative tool measures instructor proficiency and course structure (Table 3).

Table 3. *Interaction evaluation vs. OHSU course evaluation*

Interaction-Centered Course Evaluation	**OHSU course evaluation**
Student-student interaction	
Student-instructor interaction	Instructor proficiency
Student-learning material interaction	
Student-instruction interaction	Course structure

Some evaluation tools for distance education do not measure many types of student interaction. The Michigan Virtual University (2003) Online Course Evaluation Tool, for example, measures course infrastructure, technology and some aspects of student-content interaction, but does not have items for student-student, and student-instructor interactions.

Perhaps the tool used most widely for distance education evaluation is the Flashlight Current Student Inventory (CSI). The Flashlight Group was started in 1994 to address questions about technology and media use in distance education, and funded through the

US Department of Education and the Corporation for Public Broadcasting (Zuniga and Derbyshire, 2003). The CSI is a set of some 2000 questions about technology use in distance education. The CSI also contains some evaluation tools produced by Flashlight-subscribing institutions for their own studies. These tools allow educators to design evaluation studies for their own courses. Questions in the CSI have been developed and validated in association with Purdue University; they are intended to evaluate technology use in distance education (Brown, 1998), and answer questions such as:

- How do students use a particular technology?
- How is student time divided between various media (reading online, reading offline texts, viewing Web-based lectures)?
- Do multimedia methods help students learn and organize complex tasks?

Using the Flashlight CSI tools, instructors are able to characterize the contexts in which technology is used in distance education, and examine the effectiveness of that technology.

Several factors differentiate Flashlight evaluation studies from the Interaction-centered tool produced in this project (Table 4). Besides the focus of the evaluation tools (technology use vs. interaction in distance education), Flashlight tools, for the most part are developed *de novo* for individual studies (Ehrmann and Zuniga, 1997). Though validity has been established for individual questions, validity and reliability may not established all combinations and uses of these items (Hulley, 1996).

Table 4. Interaction evaluation vs. Flashlight evaluation tools

Interaction-Centered Course Evaluation	**Flashlight™ evaluation tools**
Free	$525 - $4900 per year
Ready to use	Individual evaluations must be constructed
Targeted in scope (measures students' perception of interactions)	Comprehensive group of surveys targeted at technology use in distance education
Demonstrated validity and reliability	Validity and reliability need to be determined for evaluations using CSI tools.
Interaction framework	Technology-uses framework

MATERIALS AND METHODS

Although this thesis project had materials, study methods and results, these categories do not completely describe the work accomplished. The sections in the *Materials and Methods* chapter describe the development of an interaction-centered questionnaire. The sections in the Results chapter describe the practical application of the instrument.

Objective

The purpose of this project and study has been the development and validation of an interaction-centered education evaluation tool. This process occurred in three steps:

1. Developing questionnaire items
2. Measurement study to determine the questionnaire's validity and reliability
3. Analysis of measurement study data

Each step in this process is described below.

Step 1: Developing questionnaire items

Based on review of the literature and informal discussions with distance-education faculty at OHSU, an initial set of 71 items was produced (Appendix 1). These operationalized question items were intended to measure testable qualities of findings and recommendations from the distance education literature. Faculty members evaluated these items for relevance, appropriateness, readability and simplicity. Refinements to the item set resulted in 50-item questionnaire tool, grouped into the four Moore-Hirumi categories (Appendix 2).

To test the content validity of the questionnaire, four expert reviewers rated the questionnaire items. These experts had experience in distance education and doctorate-level training in evaluation methods. Items were evaluated for content appropriateness and information value (Cox, 1996). The reviewers used a Web-based form to rate individual items and to provide feedback.

From the item ratings and expert feedback, a 32 scale-item questionnaire was created for a measurement study (Table 5). The questionnaire was implemented on the Survey Monkey Web site (www.surveymonkey.com). To improve the efficiency of the tool's administration, conditional logic was included in the Web-based tool. This logic allowed certain questions to be omitted, based on earlier responses. Most items had response 5-point Likert scale response options (range: 1 = strongly disagree to 5 = strongly agree). Categorical (ordinal) items had response numerical-range options such as *"never," "once," "two to five"* and *"more than five."*

Table 5. *Measurement study questionnaire*

Item Code	Item Description	Response Category
	Student-Classmate interaction:	
S1	I enjoyed interacting with classmates.	Likert
S2	I got to know classmates personally.	Likert
S3	I worked with fellow students on class assignments.	Categorical (ordinal)
DX	I participated in class discussions.	Categorical (ordinal)
S5	Classmates gave me feedback on my work.	Categorical (ordinal)
S6	I learned through participating in class discussions	Likert
S7	I enjoyed participating in class discussions.	Likert
SX	I interacted with classmates.	Categorical (ordinal)
TM	My interaction with the instructor occurred through:	Email
		Class discussion board
		Phone
		In person
		Other (please specify)

	Student-instructor interaction:	
T1	The instructor valued my class contributions.	Likert
T2	The instructor added valuable insights to class discussions	Likert
T3	The instructor kept class discussions on the right track.	Likert
T4	The instructor answered my questions in a timely fashion.	Likert
T5	The instructor was accessible.	Likert
TX	I interacted with the instructor.	Categorical (ordinal)
TM	My interaction with the instructor occurred through:	Email
		Class discussion board
		Phone
		In person
		Other (please specify)
	Student-Course Material Interaction:	
ML1	Lectures were relevant to course objectives.	Likert
ML2	Course lectures were easy to access/view.	Likert
ML3	Course lectures were easy to understand.	Likert
LX	I viewed the course lectures.	Categorical (ordinal)
ML5	I enjoyed watching the course lectures.	Likert
MR7	I enjoyed the assigned reading material.	Likert
ML8	I learned a lot by watching the course lectures.	Likert
MR9	I learned by reading assigned materials.	Likert
ML10	My preference for the lecture format is (or would have been):	Video with PowerPoint slides
		Audio with PowerPoint slides
		Video with hard-copy notes
		Audio with hard-copy notes
		Other (please specify)
	Student-Instruction Interaction:	
CK1	I interacted with technical support.	Categorical (ordinal)
CK2	I am satisfied with the technical support I received.	Likert
CK3	The technical support staffed answered my questions	
	In a timely fashion.	Likert
CS4	I Knew how well I was doing throughout the course.	Likert
CS5	I saw real uses for what I was learning.	Likert
CS6	I was able to choose learning materials that worked well for me.	Likert
CS7	I felt challenged by the course.	Likert

Items DX, SX, TX and LX indicate quantity-of-interaction – amount of student interaction with discussion boards, other students, instructors and online lectures. Items SM and TM indicates the type media used for interaction (email, discussion boards, phone or in person).

In addition to the 32 scale items, two global measures of perceived interaction quality were added each scale to act as dependent variables (Table 6). These items were modified versions of the global items used in Doll and Torkzadah's (1988) study of end-user computing satisfaction and were intended to be intuitive representations of student satisfaction and perceived benefit. For example, the student-student interaction scale included the global measures: *"I am satisfied with my interactions with students in the class,"* (GS1), and *"I benefited from interactions with students in the class,"* (GS2). Questionnaire item scores in each interaction scale were expected to correlate strongly with these global quality measures (Doll and Torkzadeh, 1988).

Table 6. *Global-criterion items*

Item Code	Item Description	Response Category
	Student-instruction (overall course) interaction:	
CG1	I'm satisfied with the class, overall.	Likert
CG2	The course was successful.	Likert
	Student-student interaction:	
SG1	I am satisfied with my interactions with students in this class.	Likert
SG2	I benefited from interactions with students in the class.	Likert
	Student-instructor interaction:	
TG1	I am satisfied with my interactions with the instructor.	Likert
TG2	I benefited from my interactions with the instructor.	Likert
	Student-content interaction:	
MG1	I'm satisfied with course learning materials.	Likert
MG2	I learned through interacting with course materials.	Likert

The survey included three demographic items, intended to identify students' academic backgrounds, computer experience and experience distance courses. Four open-ended questions were also included to elicit suggestions for course improvement in each of the interaction categories (Table 7).

Table 7. Demographic, Open-Ended Questions

Item Code	Item Description	Response Category
	Student demographic items:	
DE1	How many distance courses have you taken previously?	Categorical (ordinal)
DE2	What is your academic background?	Categorical
DE3	What is your computing experience?	Categorical (ordinal)
	Open-ended question:	
OP1	Suggestions on improving interactions with classmates?	Open-ended text
OP2	Suggestions on improving interactions with the instructor?	Open-ended text
OP3	Suggestions on improving course learning materials?	Open-ended text
OP4	Suggestions on improving the course, overall?	Open-ended text

Step 2: Measurement study

Participants

Participants in the measurement study were students in OHSU's Medical Informatics distance courses. Students were enrolled in either a graduate degree program or an informatics certificate program. Students were mostly working professionals in healthcare and/or information technology and had at least an undergraduate degree. All students (n = 69) in winter-term (2003) distance informatics classes were invited to participate in study. Students enrolled in more than one course were only invited to participate in one class survey. No incentives were offered.

Conducting student surveys

Fifty-five students (79% of the winter-term students) completed the online questionnaire. Survey responses were downloaded from the Survey Monkey Web site and imported into SPSS (www.spss.com) for analysis.

A convenience sample of six students was contacted for interview after their completion of the online survey. These semi-structured interviews were conducted by phone or face-to-face. User comments were recorded in written notes. Interview questions were directed to confirm that students understood the instrument's directions and to identify problems with the question items (Doll and Torkzadeh, 1888). These interviews were also included to demonstrate that the students' verbal responses generally corresponded with their questionnaire responses. Students were questioned about their overall satisfaction level and perceived benefit (i.e., the two global-criterion questions for each interaction category).

To determine the instrument's test-retest reliability, a convenience sample of nine additional students was contacted by phone or in person to complete a brief post-test questionnaire. This questionnaire was an eight-item instrument that contained the two global-criterion questions for each interaction scale. All nine students completed this questionnaire. Responses were recorded in written notes and entered into SPSS.

***Step 3:* Analysis of Measurement Study Data**

The measurement study instrument consisted of four scales. Each was intended to measure student perception of quality in one of the interaction categories. The measurement study, therefore, was intended to:

- Show that each scale was measuring what it was attempting to measure (scale validity).

- Demonstrate that items within each scale were measuring the same concept (scale reliability).

Item selection and removal

Item selection and removal was accomplished in a three-part process. First, an item-correlation matrix was constructed for each interaction scale (Appendix 3), and a corrected item-total correlation calculated for each item (Appendix 4). A corrected item-total correlation represents a test of whether one item in a scale is measuring the same concept as the other items. The corrected item-total correlation, therefore, determines if an item should be combined with other items in the scale (Aday, 1996). This value is the correlation (Pearson r) of an item with the sum of other items in the scale.

Second, items with corrected item-total correlations $r < .40$ were removed from each scale (Table 8). This cutoff value was chosen because it was used in the construction of the 36-item Health Status Questionnaire used in the 1994 Medical Outcomes Study. A corrected item-total correlation ($r \geq .40$) was judged to signify a significant and robust association between items (Aday, 1996). Question items remaining had corrected item-total correlations of $r \geq .40$, significant at $p < .05$.

Third, as a determinate of criterion-related validity (Hulley, 2001), the correlation of each item with the sum of the two global-criterion items (for that scale) was calculated for each scale (Appendix 3). The global-criterion question items were assumed to be valid measures of satisfaction and perceived benefit (Doll and Torkzadeh, 1888). The item/global-criterion correlation, therefore, was used as a measure of criterion-related validity. Items with item/global-criterion correlations $r < .40$ were removed from each scale (Table 8). This cut-off point was arbitrary, as there is no acceptable standard for item/global-criterion correlation (Doll and Torkzadeh, 1888). Items remaining in each scale had item/global-criterion correlations of $r \geq .40$, significant at $p < .05$.

Table 8. *Items Removed from Interaction-Centered Evaluation Tool*

Item Code	Item Description	Removal Criteria
	Student-classmate interaction scale questions:	
S2	I got to know classmates personally.	Corrected item-total correlation, Item-global r < .40
S3	I worked with fellow students on class assignments.	Corrected item-total correlation, Item-global r < .40
DX	I participated in class discussions.	Corrected item-total correlation, Item-global r < .40
SX	I interacted with classmates.	Corrected item-total correlation, Item-global r < .40
	Student-instructor interaction scale question:	
TX	I interacted with the instructor.	Corrected item-total correlation, Item-global r < .40
	Student-learning material interaction scale questions:	
LX	I viewed the course lectures.	Item-global r < .40
ML10	My preference for the lecture format is (or would have been):	Not a scale item
	Student-instruction interaction scale questions:	
CK1	I interacted with technical support.	Item-global r < .40
CK2	I am satisfied with the technical support I received.	Item-global r < .40
CK3	The technical support staffed answered my questions	Item-global r < .40
	Demographic questions:	
DE1	How many distance courses have you taken previously?	Chi Square p > .05
DE2	What is your academic background?	Chi Square p > .05
DE3	What is your computing experience?	Chi Square p > .05
	Type-of-interaction questions:	
SM	My interaction with the other students occurred through:	Chi Square p > .05
TM	My interaction with the instructor occurred through:	Chi Square p > .05

The result of this process was the removal of 15 items from the questionnaire. Two items in the student-classmate scale relating to student collaboration, "*I got to know classmates personally*" (S2) and "*I worked with fellow students on class assignments*"(S3) were removed because of low corrected item-total and item/global-criterion correlation. The four frequency-of-interaction items (DX, SX, TX, LX) and the technical-support interaction items (CK1, CK2, CK3) were also removed as scale items because of low corrected item-total and item/global-criterion correlation (Appendix 5A). These frequency-of-interaction items are included in the final instrument as non-scale items, and are not part of the reliability calculations.

Nonparametric tests were performed on the three demographic items to determine if difference existed between groups for the medians of any of the eight global-criterion variables. Nonparametric tests were used because the distribution of the groups (e.g., students who had taken no previous distance courses vs. students who had taken one previous course vs. students who had taken several previous courses) was unknown. Median tests were performed on items "*How many distance courses have you taken previously?*" (DE1), "*What is your academic background?* (DE2) and "*What is your computing experience?*" (DE3). None of these tests demonstrated a significant difference (Chi Square $p < .05$) between groups for the eight criterion variables, and the three demographic items were removed from the instrument (Appendix 5B).

Median tests were also performed on the types-of-interaction variables "*My interaction with other students occurred through*" (SM) and "*My interaction with the instructor* *occurred through*" (TM), to determine if the method of communication (email, discussion boards, phone or in person) had bearing on course satisfaction or perceived benefit. Interestingly, students who reported interaction with the instructor(s) through class discussion boards or in person, had significantly higher levels of course satisfaction than those students who reported no interactions in these categories. The other global variables, including satisfaction with student-instructor interactions, however, were not significantly higher for these groups. None of the other tests demonstrated a significant difference (Chi Square $p < .05$) and both type-of-interaction questions were removed from the instrument (Appendix 5B).

The resulting 36-item Interaction-Centered Course Evaluation (ICCE, Table 9) was divided into 20 scale items, eight global-criterion items, four frequency-of-interaction questions, and four open-ended questions.

Table 9. The Interaction-Centered Course Evaluation (ICCE)

Item Code	Item Description	Response Category
	Student-Classmate interaction:	
SG1	I am satisfied with my interactions with students in this class.	Likert
SG2	I benefited from interactions with students in the class.	Likert
S1	I enjoyed interacting with classmates.	Likert
S5	Classmates gave me feedback on my work.	Categorical (ordinal)
S6	I learned through participating in class discussions	Likert
S7	I enjoyed participating in class discussions.	Likert
	Student-instructor interaction:	
TG1	I am satisfied with my interactions with the instructor.	Likert
TG2	I benefited from my interactions with the instructor.	Likert
T1	The instructor valued my class contributions.	Likert
T2	The instructor added valuable insights to class discussions	Likert
T3	The instructor kept class discussions on the right track.	Likert
T4	The instructor answered my questions in a timely fashion.	Likert
T5	The instructor was accessible.	Likert
	Student-Course Material Interaction:	
MG1	I'm satisfied with course learning materials.	Likert
MG2	I learned through interacting with course materials.	Likert
ML1	Lectures were relevant to course objectives.	Likert
ML2	Course lectures were easy to access/view.	Likert
ML3	Course lectures were easy to understand.	Likert
ML5	I enjoyed watching the course lectures.	Likert
MR7	I enjoyed the assigned reading material.	Likert
ML8	I learned a lot by watching the course lectures.	Likert
MR9	I learned by reading assigned materials.	Likert
	Student-Instruction Interaction:	
CG1	I'm satisfied with the class, overall.	Likert
CG2	The course was successful.	Likert
CS4	I knew how well I was doing throughout the course.	Likert
CS5	I saw real uses for what I was learning.	Likert
CS6	I was able to choose learning materials that worked well for me.	Likert
CS7	I felt challenged by the course.	Likert
	Open-ended and Frequency of Interaction:	
OP1	Suggestions for improving interactions with classmates?	Open-ended text
OP2	Suggestions for improving interactions with the instructor?	Open-ended text
OP3	Suggestions for improving course learning materials?	Open-ended text
OP4	Suggestions for improving the course, overall?	Open-ended text
SX	I interacted with classmates.	Categorical (ordinal)
TX	I interacted with the instructor.	Categorical (ordinal)
LX	I viewed the course lectures.	Categorical (ordinal)
DX	I participated in class discussions.	Categorical (ordinal)

Internal-constancy reliability

After item selection/removal, the ICCE instrument was evaluated to determine its reliability. As a determinant of internal-consistency reliability (Aday, 1996), Chronbach's *alpha* was calculated for items in each scale, and for all scale items combined. The *alpha* value is a measure of the correspondence between answers to sets of items in a scale. For example, responses to half the items in a scale could (arbitrarily) be selected and correlated with the responses to the other half. This would serve as a measure of the variation (reliability) of the different sets of responses. *Alpha* is the calculation of set-response correlations between every possible combination of item sets. For a given scale, *alpha* values of .70 are considered acceptable, values of .90 considered optimal (Aday, 1996).

For student-classmate, student-instructor, student-learning material and student-instruction interaction scales, *alpha* values were .89, .93, .92, and .86, respectively. *Alpha* for the four scales combined was .95. These results demonstrate the internal-consistency reliability of the four interaction scales and for the instrument as a whole.

Test-retest reliability

Spearman rank-order coefficients (*rho*) for the test- and retest-scores were calculated for the eight global-criterion items (Appendix 6). The size of the retest sample was small (n = 9) and test-retest reliability (Aday, 1996), for the instrument was not demonstrated: Only 3 of the eight global-criterion items showed a significant coefficient value at $p < .05$ (two-tailed).

RESULTS

The 36-question ICCE may be used summatively or formatively to evaluate distance-education courses and course components, in the following ways:

Summative evaluation

- No standard for item or scale scores was produced in this study. It is reasonable, however, to assert that if, for a particular course, the mean for the item *"I was satisfied with the course, overall"* corresponds to a (Likert-scale) response of *"I agree,"* and if the scales are valid and reliable, then students are, in general, satisfied with the course. Therefore, it is possible to set some arbitrary benchmarks for the tool: Individual item scores ≤ 3.5 on Likert-scale may indicate areas ripe for improvement and scores ≤ 3.0 on Likert-scale questions may indicate possible problem areas.
- Individual item scores and interaction scale scores can be compared across courses.
- Individual item scores and interaction scale scores for courses can be followed across time to gauge progress or evaluate interventions.
- The two global student-instruction items (CG1 and CG2, Table 6) can be used to approximate student satisfaction and perceived benefit for the course.

Formative evaluation

- Scores may identify opportunities for intervention. For a particular course, a mean item score ≤ 3.5 may identify an area ripe for improvement.

- Individual item scores can be followed across time to gauge progress or evaluate interventions.
- Examination of individual item and scale evaluation scores across a group of courses may identify or exclude potential targets for technology improvements (e.g. online lectures, Web-based discussion boards).

The cases below present course-specific data and recommendations from the measurement study. In these cases, some of the identifying information (such as course names) has been changed. As discussed above, the ICCE is not a standardized instrument, and designation of an acceptable versus a substandard score is arbitrary. For the purpose of these hypothetical exercises an item with mean Likert score of ≤ 3.5 will represent a target for scrutiny. An item with mean Likert score ≤ 3.0 will represent a potential problem area. Informal qualitative content analysis was used to identify themes within the student comments generated by the open-ended questions.

Case 1: Biometry

Biometry Survey was an existing on-campus course developed into distance format, and offered by distance for the first time in the winter term of 2003. A lecture format new to the informatics distance program – video synchronized with PowerPoint slides – was also being implemented for the first time. The instructor would like to know if the course and the new lecture format work for the students.

Case 1 survey data:

Course mean scores for items in all scales were > 3.5 except for the T2 item in the student-instructor interaction scale are shown in Table 10 and Figure 3.

***Table 10.** Student-instructor interaction mean scores*

Item Code	Item Description	Mean Score	Score Category
TG1	I am satisfied with my interactions with the instructor.	4.0	Likert
TG2	I benefited from interactions with the instructor.	4.0	Likert
T1	The instructor valued my contributions.	3.8	Likert
T2	**The instructor kept class discussions on the right track.**	**3.5**	Likert
T3	The instructor added valuable insights to class discussions.	3.8	Likert
T4	The instructor answered questions in a timely fashion.	4.0	Likert
T5	The instructor was accessible.	4.3	Likert

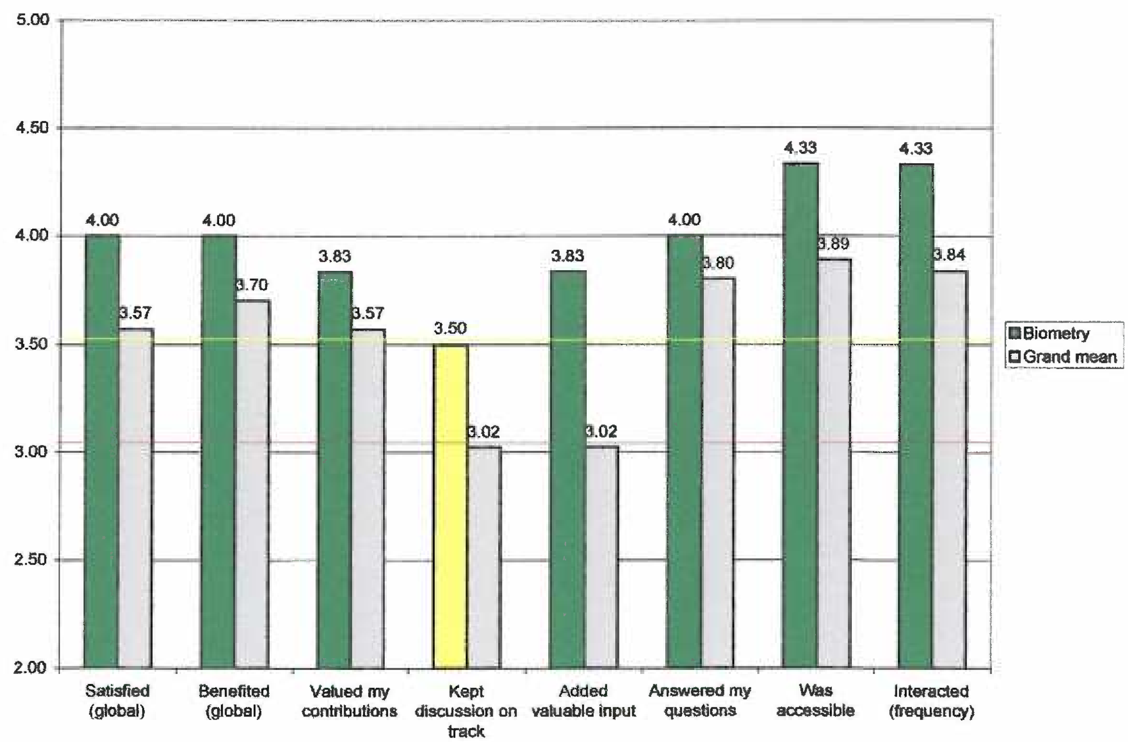


***Figure 3**. Student-instructor interaction mean Likert scores – Grand mean scores (grey bars) represent mean scores for that item of all students in all seven (Winter term, 2003) informatics distance classes; yellow or red bars represent scores below benchmark.*

Case 1 Student comment themes:

1. **Students would have liked more instructor guidance and participation in the discussion boards:**

   "I think it would help if there are more guided discussions. Sometimes I am not sure what aspect to discuss."

   "More frequent responses from the instructor on the discussion board."

   "The instructors entered into the discussion board discussions occasionally. The instructors in other courses entered the discussions more frequently (it seems to me) and often posed additional questions that fed off the current discussion thread. That seemed to keep the momentum going as the students added their own thoughts to the thread. Sometimes quite involved discussions or tangents developed."

2. **Students were generally pleased with course learning materials:**

   "The text (Pagano) was one of the best statistics books I've read. The author repeated concepts and explanations throughout the text and didn't assume that you had 'memorized' the book up to that point when he introduced a problem. The solutions to the problems in the text were thoroughly discussed."

   "If it hadn't been for Dr. Kraemer's lectures on 'how to' use SPSS with step-by-step instructions for setting up the calculations I don't think we could have used the program effectively. Even though the course is over I've taken those notes and put them with my SPSS manual for future reference."

   "I like the exercise a lot. I think it will help also if there are more examples (maybe outside the video) with doing analysis with SPSS plus giving us more insight on how to interpret the results."

   However:

   "For using slow connections (such as when traveling) it would be helpful to have an audio plus printed out notes option to make it easier use RealPlayer."

   "Break the lectures into shorter segments. Many of them were about an hour, which is too long. 20-30 minutes would be better. More examples on using SPSS and how to interpret the results SPSS-generated tables."

3. **Fewer homework problems should be assigned:**

"Lots of homework which seemed out of proportion unit wise . . ."

"Having fewer problems to do each week might have allowed more time for absorbing the material or entering into discussions on the discussion board. It is a difficult issue however because how do you learn to solve problems unless you do a bunch of them? On the other hand for a course billed as a 'survey' maybe we don't need to do twenty plus calculations each week. For me the jury is still out on that issue. In a few months I may think differently ..."

"Have more analysis results interpretation and perhaps real life application example on how these analyses are being used."

"Fewer homework problems. Doing the homework assignments took an inordinate amount of time. More extensive feedback on homework assignments. I would have liked to have the assignments sent back to me either via e-mail or U.S. mail with suggestions and criticism."

Case 1 discussion:

The biometry course scored well in all interaction categories. Only one scale item signaled a possible area for course improvement. The "*instructor kept course discussions on track*" (T2) item scored below the benchmark. This item corresponded to one of the themes in the student comments: *students would have preferred more instructor participation in the discussion groups.* However, the goal of the instructor was to let students try to answer discussion board questions before instructor intervention. Therefore, this theme is not necessarily inconsistent with the course plan.

High scores for learning-material interaction represent a good level of satisfaction with course lectures. These lectures were perceived as relevant, easy to view, easy to understand and enjoyable. The students also perceived they were learning from the lectures. Similarly, the students enjoyed and learned from course reading material (Table

11, Figure 4). These scores are reinforced in the positive student comments about the course learning materials.

Table 11. Student-learning material interaction mean scores

ITEM CODE	ITEM DESCRIPTION	MEAN SCORE	SCORE CATEGORY
MG1	I'm satisfied with course learning materials.	4.0	Likert
MG2	I learned through interacting with course materials.	4.5	Likert
ML1	Lectures were relevant to course objectives.	4.7	Likert
ML2	Course lectures were easy to access/view.	4.3	Likert
ML3	Course lectures were easy to understand.	4.2	Likert
ML5	I enjoyed watching the course lectures.	4.2	Likert
MR7	I enjoyed the assigned reading material.	4.2	Likert
ML8	I learned a lot by watching the course lectures.	4.7	Likert
MR9	I learned by reading assigned materials.	4.5	Likert

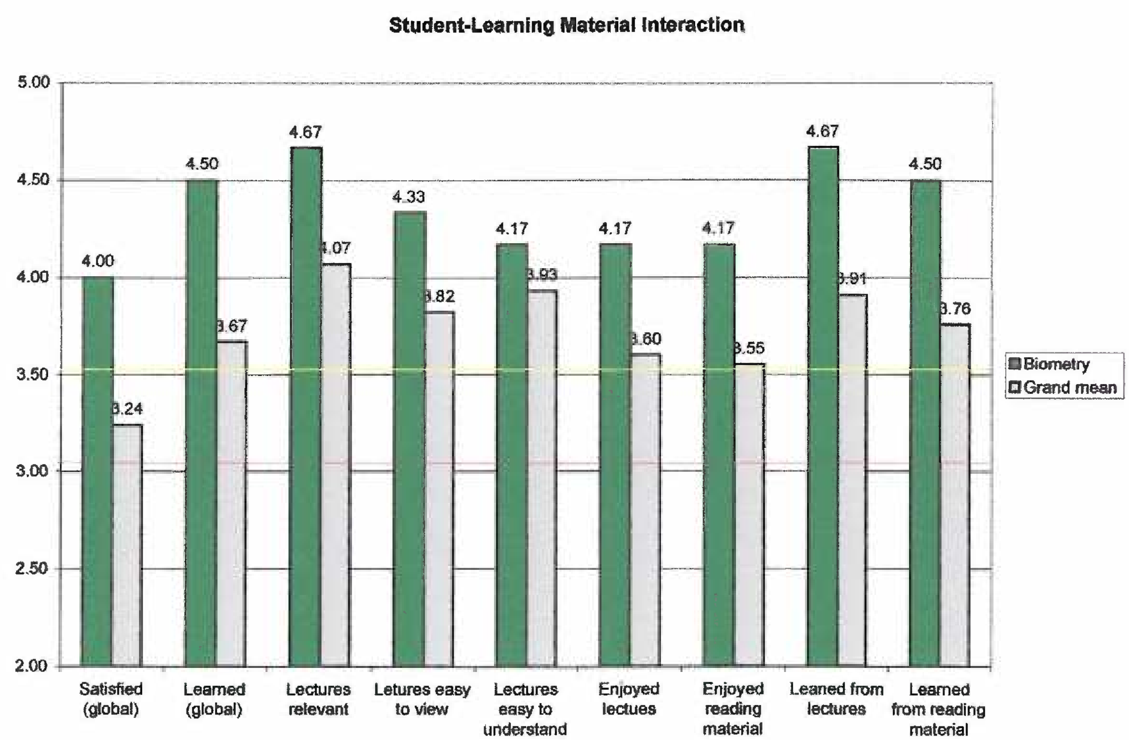


***Figure 4**. Student-learning material interaction mean scores – Grand mean scores (grey bars) represent mean scores for that item of all students in all seven (Winter term, 2003) informatics distance classes.*

Case 1 analysis/recommendations:

1. As demonstrated by the high student-instruction interaction mean scores (CG1 = 4.2, CG2 = 4.3, Figure 5), the course is a perceived success. The high global scores in the three other interaction scales reinforce this conclusion.
2. Students would appreciate more instructor input and guidance on class online discussions.
3. High scores for learning-material interaction suggest that the new lecture format (video with synchronized slides) works well.
4. The course text works well.
5. Students would appreciate fewer homework problems.

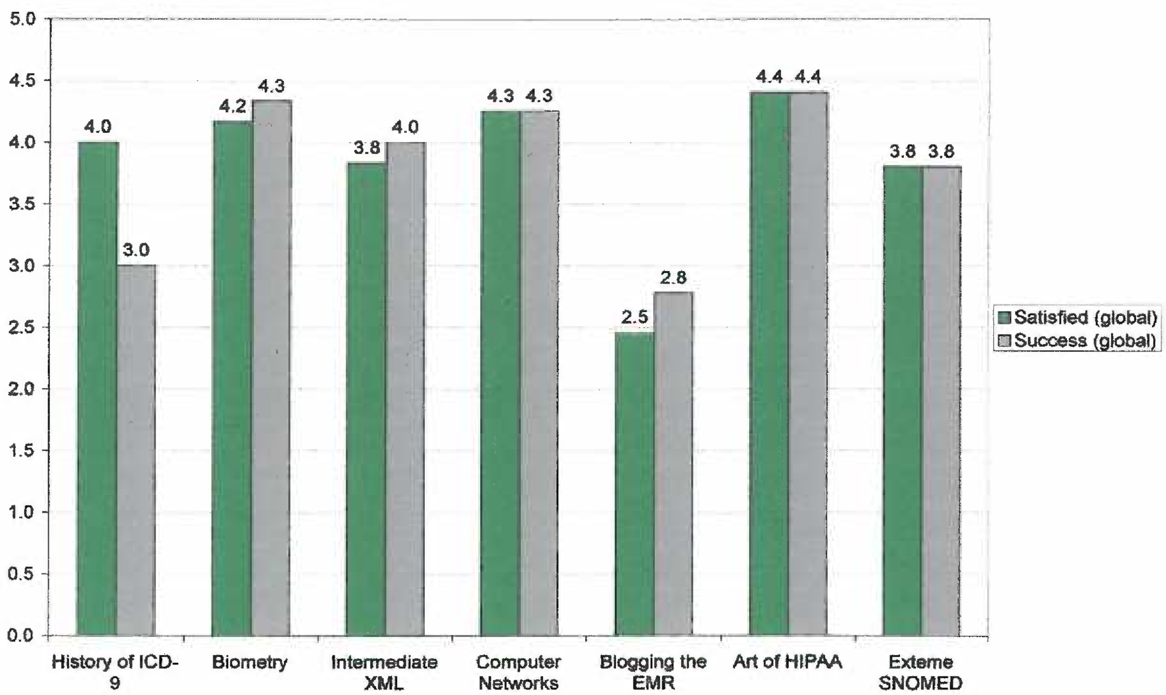


Figure 5. Global-criterion variable scores for all courses – *Grand mean scores (grey bars) represent mean scores for that item of all students in all seven (Winter term, 2003) informatics distance classes.*

Case 2: Intermediate XML

Intermediate XML (not the real course title) was taught for the first time by distance during the winter term of 2003. An instructor remote to the OHSU campus taught the course by distance. Audio lectures used in the course had been recorded from an on-campus course in the year 2001. Other learning materials (such as class lecture notes) had also been developed for the on-campus class. On-campus students as well as distance students took the course as a Web-based (distance) course. The instructor and program administrators would like to know how the student-student and student-instructor interactions worked under these circumstances. They would also like to know what problems (if any) were encountered with the learning materials, and if there are specific targets for course improvement.

Case 2 survey data:

Overall satisfaction with classmate interaction (SG1) was below the 3.5 benchmark.

Other items in the classmate scale were above (Figure 6).

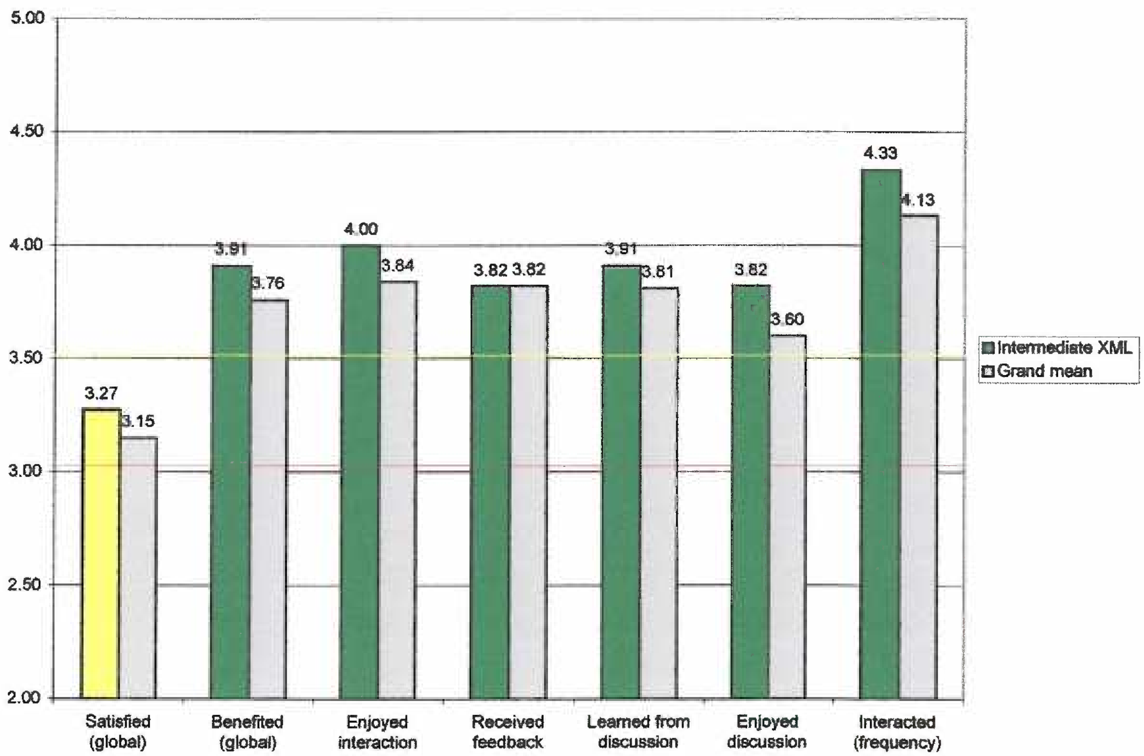


Figure 6*. Student-classmate interaction mean Likert scores for Intermediate XML –* Grand mean scores (grey bars) represent mean scores for that item of all students in all seven (Winter term, 2003) informatics distance classes*; yellow or red bars represent scores below benchmark.*

Three items in the student-instructor scale fell below the 3.5 benchmark: *"value of contributions"* (T1), *"kept discussion on track"* (T2), and *"answered questions in a timely fashion"* (T3) (Figure 7).

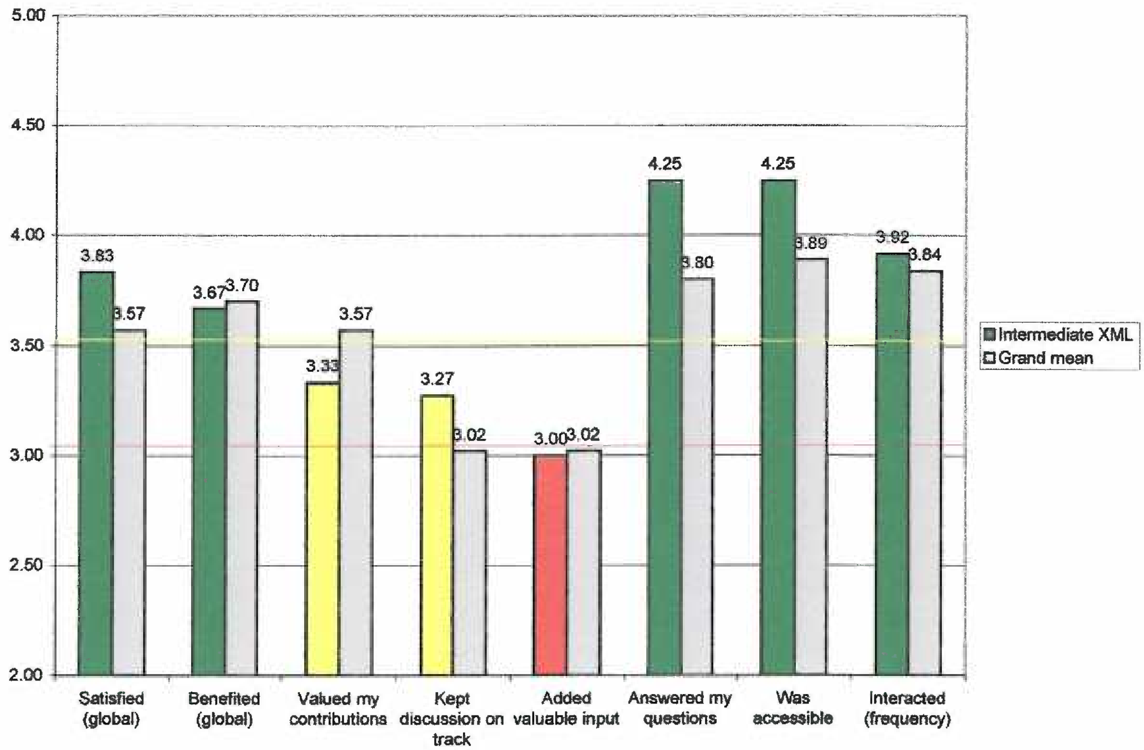


Figure 7*. Student-instructor interaction mean Likert scores for Intermediate XML –* Grand mean scores (grey bars) represent mean scores for that item of all students in all seven (Winter term, 2003) informatics distance classes*; yellow or red bars represent scores below benchmark.*

Three items in the student-learning materials interaction scale also fell below the (3.5) benchmark: *"satisfaction with learning materials"* (MG1), *"enjoyed lectures"* (ML5), and "*enjoyed reading material*" (MR7) (Figure 8).

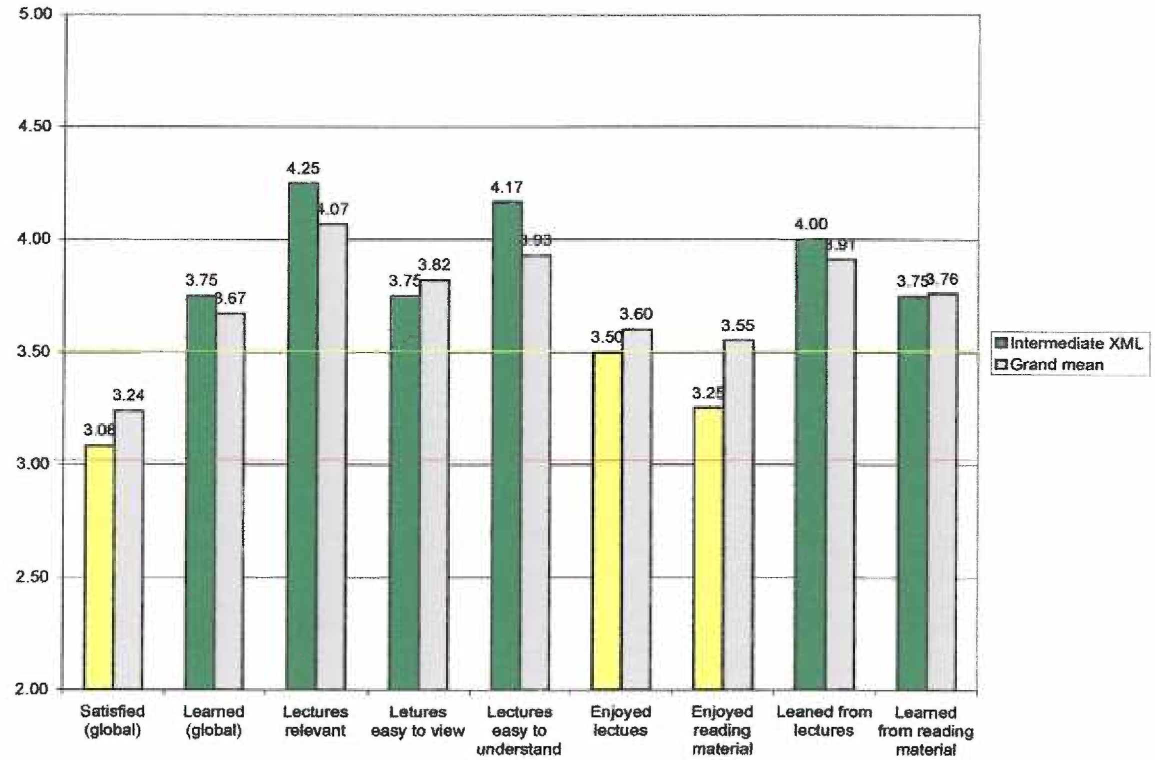


Figure 8. *Student-learning material interaction mean Likert scores for Intermediate XML – Grand mean scores (grey bars) represent mean scores for that item of all students in all seven (Winter term, 2003) informatics distance classes; yellow or red bars represent scores below benchmark.*

Students gave low marks to two items in the student-instruction interaction scale: knowledge of how they were doing during the class (CS4) and ability to choose learning materials that worked for them (CS6) (Figure 9).

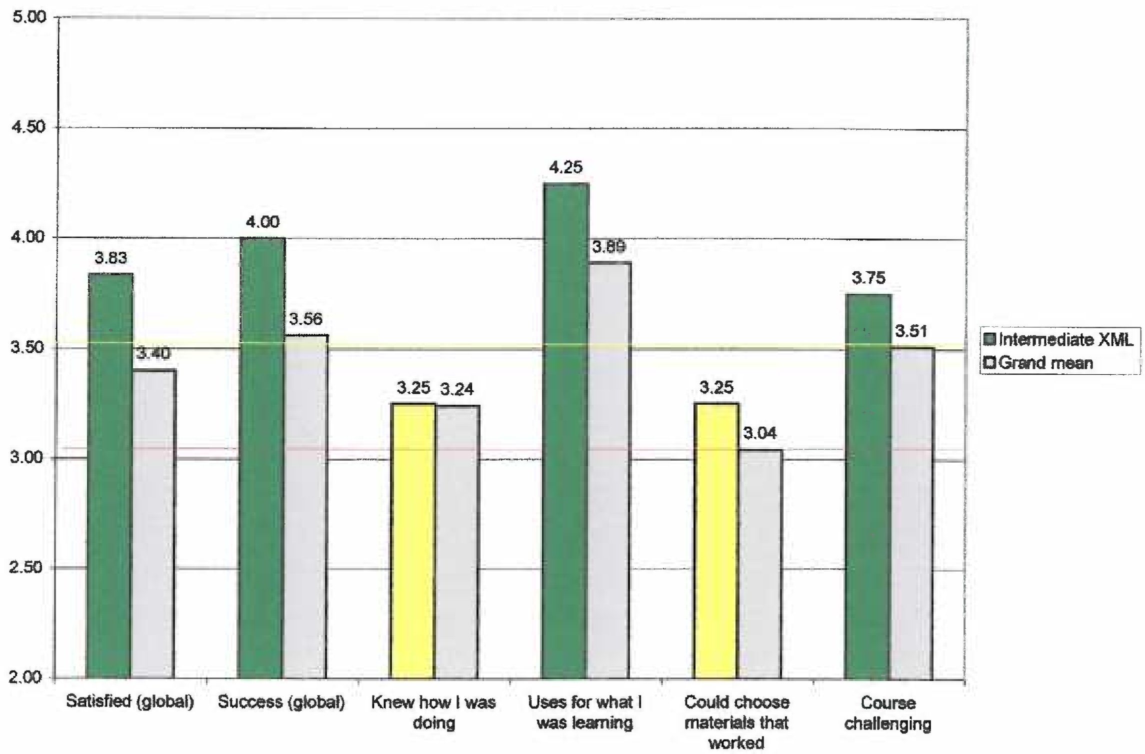


Figure 9. *Student-instruction interaction mean Likert scores for Intermediate XML –* Grand mean scores (grey bars) represent mean scores for that item of all students in all seven (Winter term, 2003) informatics distance classes; *yellow or red bars represent scores below benchmark.*

Case 2 student comment themes:

1. **The discussion boards were large and difficult to use:**

   "I think this [course] discussion board was a bit more complicated since the class size was so large compared with my previous classes."

   "I don't think it's technically feasible but the nature of the discussion board makes it hard to look for a previous posting [...] More robust discussion board features would help interactions."

   "I found it rather cumbersome to go through the large number of postings each week. Perhaps it would be better to break people into groups. That way there would be more personal interaction and the discussion boards would not seem so unwieldy."

   "People posted to the discussion site but then did not thread as frequently as other distance courses. It seems like they wanted to post their answer but were not interested in interacting with other students."

2. **The class was somewhat impersonal:**

   "I hadn't really thought about it but I really didn't get to know my classmates much at all."

   "With distance learning it is difficult to get to know one's classmates unless the students interact with each other outside of the course. As noted by fellow students and recognized by the instructor with so many students in the class at times the discussion seemed overwhelming."

3. **Student preferred more instructor input on course discussions:**

   "It would have helped if the instructor had added more to the discussion boards."

   "There also seemed to be less instructor comments when compared to [other] classes and this may have been due to the size of the class again. I thought the instructor would have participated as much as in my previous classes but it wasn't the case in this class. I mainly learned from the other students in the class."

   "The instructor or a teacher aide should be screening the postings for content rather then just counting them for participation."

   "I'd like to see more instructor comments on the discussion board as in previous courses in the distance program."

One student would have preferred more direct interaction with the instructor: "I believe I would have benefited from direct communication with the instructor regarding my personal situation and career path."

Another commented:
"It seemed that the instructor played the role of coordinator and that interactions with the instructor were only necessary for items pertaining to course infrastructure rather than content. I would like to see the instructor for this course play a more instructor-like role."

4. **Problems with course learning material and media:**

"Would like to have video when [the] lecturer points on the diagrams what he/she is talking about."

"Quality of handouts was often difficult to read. Sound quality of lectures was often poor."

"Would like to see different textbook. I felt that often a comment would be made in the lecture that differed from that in the book."

"For slides that are difficult to read one could put fewer slides on the page or make the PowerPoint slide available for local printing (did occur on occasion in this course and more frequently in another course I had taken)."

"1. Put the PowerPoints on a standard format – 6 slides to a PDF page – for easier printing. 2. Put the audio into a downloadable mp3 format for CD burning or else post instructions on how to use programs such as Total Recorder ($12) to capture streamed audio into a recordable format. 3. Test every lecture for listenability by hooking up to a 56Kbs dial-up and listening to the streamed audio for clarity. On the OHSU side, make sure the microphone and recording quality are acceptable."

5. **Quality of lecture content:**

"Some of the lectures were older and were not quite up-to-date."

"Lecturers follow order of bullets on slide but not just read them. More real-life examples in correlation with topics."

"Often felt that the materials were of a retread nature."

6. **Quizzes, exams:**

"The quiz material should be based more on the lectures. Some of the quiz material was difficult to match to the appropriate reference."

"Quizzes often focused on small details rather than important concepts."

"Participation grades should be posted weekly not at the end of the course. We only had quiz score to match our grades."

Case 2 Discussion:

One theme that emerged from comments on student-instructor interaction was preference for more instructor involvement on class discussion boards. This theme is consistent with the low scores on the student-instructor interaction scale, especially items *"The instructor kept class discussions on the right track"* (T2), and *"The instructor added valuable insights to class discussions"* (T3).

Another theme in the student comments was the size and unwieldiness of the discussion boards. This could be contributing to low score for overall satisfaction of student interaction.

The many suggestions within the student comments for improvement in course learning materials and media underscore the learning-material interaction scale responses. The students felt that they were more able to learn from the learning materials (MG2, ML8, MR9 > 3.5) than they enjoyed, or were satisfied with these materials (MG1, ML5, ML7 ≤ 3.5). Student comments single out problems with the content of online lectures, such as *"not quite up-to-date," "retread,"* and *"needing more real-life examples."*

The low score on item *"I knew how well I was doing throughout the course"* (CS6) is consistent with the several student suggestions regarding the appropriateness of and feedback from quizzes.

Case 2 analysis/recommendations:

1. As demonstrated by the student-instruction interaction scores above the 3.5 benchmark, students were satisfied with the course overall and felt they benefited from the course. These scores fell in the midrange of the seven courses taught during this term (Figure 2).
2. Discussion groups should be smaller, more structured.
3. Student would prefer more instructor input and guidance on course discussions.
4. Discussion group size and lack of guidance may be contributing to feelings of student isolation.
5. Course lectures should be improved; lectures should be more interesting and enjoyable.
6. Course reading material should be improved.
7. Other course instructional materials (e.g. handouts, slides, audio) should be improved.
8. Quizzes should be more in line with course materials.
9. Students would like more feedback on the quizzes.

DISCUSSION

Several point deserve further discussion. The following sections elaborate on some of the issues in development and evaluation of this questionnaire.

Item removal

Items removed from the instrument were those with low corrected item-total and/or low item-global criterion item correlation scores. Removed items, therefore, were determined to be measuring different parameters than remaining items. In other words, items were removed because they did not fit the same interaction constructs as the remaining items.

There are other possible explanations for the low correlation scores other than that the items were measuring different constructs. For example, two items removed from the instrument, "*I got to know classmates personally*" (S2), and "*I worked with fellow students on class assignments*" (S3) were intended to measure elements of student-student collaboration. The seven courses included in this study, however, did not encourage or provide vehicles for student-student collaboration. It is likely that since there was little collaboration occurring, the two items were not considered applicable by the participating students. A study with a more diverse group of courses may justify inclusion of these items.

Another possibility for the low correlation scores is that the student sample was too small to pick up significant differences. Although none of the individual frequency-of-interaction items (DX, SX, TX and LX) met the selection criteria, the *sum* of the student interaction scores (the sum items DX, SX, TX and LX) did correlate significantly with overall course satisfaction and perception of course success ($r \geq .40, p \leq .01$) items (Appendix 7). This is an indication that the study sample may have been too small to pick up some significant correlations. For example, for a sample size of 47 and a target correlation $r \geq .40$ with *alpha* (two-tailed) = .05, the power is .80; for sample size 75 and a target correlation $r \geq .40$ with *alpha* (two-tailed) = .05, the power is .95 (Hulley, 1996). Since the sample size for the study was 55, the probability of missing a real association is about 17%. A study with a greater number of students may be able to determine whether these frequency-of-interaction questions should be included in the instrument.

How the results of this study relate to its theoretical basis

The theoretical basis for the ICCE is supported by research in distance education and recommendations of distance-education consortia (Table 1). Factors that improve student interaction should improve the course, and consequently, should improve student satisfaction and perceived success. We would therefore expect significant correlation between interaction scale items and course satisfaction and perceived success.

The data from the measurement study support this relationship. Individual scale items are significantly correlated with overall course satisfaction (CG1) and perception of course success (CG2) items. Correlation of all individual scale items (except items "*Classmates gave me feedback on my work,*" and "*Course lectures were easy to access/view*") with overall course satisfaction and perceived course success items was ≥ .40 ($p \leq .01$) (Appendix 7). Total student interaction scores (the sum of all quantitative measures, items DX, SX, TX and LX) also correlated significantly with overall course satisfaction and perception of course success items ($r \geq .40$, $p \leq .01$). Two conclusions can be drawn from these correlations: (1) The better the quality of students' interactions, the more they were satisfied with the course and felt the course was a success, and (2) the more students interacted, the more they were satisfied with the course and felt the course was a success. Since it is reasonable to look at course satisfaction and perceived course success as direct measures of student benefit, it is reasonable to say that that course interactions benefit students. In other words, results of this study suggest support for the ICCE's theoretical basis: Student interaction improves courses and benefit students.

Test-retest reliability

This study has demonstrated the validity and criterion-related reliability of the ICCE instrument. The study, however, failed to demonstrate test-retest reliability for five of the eight global-criterion items. Two possible explanations for this deficit are the test-retest design and the limited number of subjects tested.

In the first case, the design of the test-retest evaluation was problematic. Students who responded to questionnaire items by Web-based questionnaire in the original test were asked the questions verbally (by telephone or in person) on the retest. This could have introduced a systematic response error on the retest. Students may have had some reticence to answer "judgmental" questions by telephone about their classmates and instructors. Scatterplot graphs for the test vs. retest scores (Appendix 8) of the two student-instructor interaction variables support this idea – all of the retest (verbal response) scores are higher than the test (Web-based response) scores.

It is also possible that the number of subjects included in the retest (n = 9) was too small to pick up significant correlations (Spearman's *rho*) between test scores and retest scores. Test-retest reliability scores of .90 or above are preferred in studies that track changes in individuals over time especially if the observed changes are due to treatment effects. In studies where the focus is on changes in scores at the group-level, reliability scores of .70 or above are acceptable (Aday, 1996). If we select a target reliability score of .80 (two-tailed *alpha* = .05, *beta* = .20), then n = 9 is an appropriate sample size (Hulley, 1996, p218). Test results in this study support this estimate: The three test-retest correlation scores ≥ .80 were significant at $p < .50$ (two-tailed), (Appendix 6). However, since test-retest correlation scores in this study attempt to demonstrate the stability of attitudes among a group of students, the .70 reliability score (two-tailed *alpha* = .05, *beta* = .20) would have been better target, and a sample size of 13 appropriate (Hulley, 1996, p218).

Criterion-related validity

Demonstration of the instrument's criterion-related validity was based on the assumption that the global-criterion variables were valid measures of student satisfaction and perceived benefit. It is a reasonable assumption that the statement *"I was satisfied with the course, overall"* will be taken at face value and interpreted correctly. This has not been demonstrated, however, through either quantitative or qualitative methods. Although we can calculate a mean score for each scale item, and be reasonably sure that it is valid and reliable, we cannot determine what that quality score means, in absolute terms. A better determinate of criterion-related validity could be accomplished through factor analysis (Kerlinger, 1978) and multitrait-multimethod analysis (Doll and Torkzadah, 1988).

Areas for further research

Several avenues of research are suggested by this work. Exploration of the survey instrument through *factor analysis* and *multitrait-multimethod analysis* would be a logical next steps in establishing the underlying constructs, and demonstrating the validity of the instrument. These methods require more subjects than were available in this measurement study. In this section, suggestions are offered for an enlarged study utilizing factor analysis and multitrait-multimethod analysis. Other areas for further research outlined here are determination of student benefit and standardization of the ICCE instrument.

Determination of underlying constructs through factor analysis and multitrait-multimethod analysis

Although the interaction constructs used in this study are rooted in established theory, the scale categories in the ICCE instrument have not been defined by statistical methods, and are therefore somewhat arbitrarily. While the *"student-student," "student-instructor," "student-content"* and *"student-instruction"* categories have intuitive meaning, they may not be the best fit for the actual underlying constructs. For example, *"collaboration," "active learning"* or *"feedback"* might be better categories. The use of scale groups with known (statistically-defined) constructs would be preferable to the use of scale groups with arbitrary-defined constructs. Statistically-defined constructs could be determined though factor analysis and multitrait-multimethod analysis

Factor analysis

Determination of constructs underling student attitudes toward course interactions would be an appropriate step in the exploration of the ICCE evaluation tool. *Factor analysis* has been widely used in surveys where attitudes, traits and cultural patterns are measured (Kerlinger, 1978). Factor analysis is a statistical method for identifying correlation relationships between variables and determining underlying groupings – that is, which variables go together. These groups, or *factors*, may be obvious or non-apparent. Variables that correlate highly with other variables in the same group are said to *load* on that factor and "belong" in that group. Once factors are established, it is up to the researcher to determine the meaning and ontological organization of these constructs.

If constructs underlying student interaction can be identified and validated, we are in a better position to explore questions regarding the components of interaction. For example, if we determine that "*quantity of student-instructor interaction*" is one of the factors, we might wish to test the hypothesis that time spent in student-instructor interaction is positively correlated with student grades. Or we might wish to test the hypothesis that students with higher bandwidth Internet connections have a greater number of student-instructor interactions than their classmates.

Multitrait-multimethod, convergent and discriminant validity

We may also wish to test the convergent and discriminant validity of the instrument through *multitrait-multimethod (MTMM) analysis* (Doll and Torkzadah, 1988). MTMM is a method for comparison of variables across factor groups. In a 30-item survey, for example, each item would be compared with all others in a 30 x 30 correlation matrix. This matrix would consist of $(n^2 - n)/2 = (435)$ different *r* correlation values. If factor groups represent valid categorizations of items, we would expect that correlation between items within factor groups would be significantly greater than zero – this is a test of convergent validity. We would also expect that correlation between items would be significantly higher than correlation between items of different factor groups – this is a test of discriminant validity.

Design of a factor analysis/MTMM study

The same Web-based survey instrument could be used in a factor analysis study. The particulars of the study population and administration procedures would be influenced by the instructors willing to allow their students to participate in the study. Factor analysis requires a large sample size to avoid over-fitting of variables. A 10:1 ratio of sample number to variable number is recommended (Doll and Torkzadah, 1988). There are 27 interaction scale items in the survey, suggesting a sample size of $n \geq 270$. The original measurement study included 55 students in 7 courses over one term. Increasing the sample size to the number required by factor analysis would require a 5-fold increase in the number of participating students.

Score standardization

No standard for item or scale scores was produced for this study. Ideally, the ICCE should produce a standardized interaction score based on a simple equation such as $I_{TOTAL} = I_S + I_T + I_C + I_N$, where I_{TOTAL} represents the total interaction score, I_S represents a student-student interaction score, and so forth. We would like the ICCE to be standardized across a wide variety of courses, subjects, students and schools.

What do students mean by "benefit?"

It would be good to know what students perceive as "beneficial" in distance education. Social constructivist theory and the distance education literature lead to the conclusion that student interaction leads to higher student satisfaction and some improved outcomes. Theories and studies do not tell what student would like, specifically, from distance education. It may seem apparent that students want to learn specific skills, make social connections, establish academic networks, get good grades and get good jobs. But how do we know this? How important is each of these factors? What other benefits would students like to derive from the education experience? A qualitative study using grounded theory and content analysis should answer these questions.

CONCLUSION

The accessibility of low-cost, high-quality courses is firing the revolution of distance education. The availability of free, easy-to-use evaluation tools should aid in the assessment and improvement of distance courses. The ICCE questionnaire developed in this project measures student perception of quality of course interactions with demonstrated validity and reliability. It is free, easy for educators to implement and easy for students to use. The Web-based nature of the tool also allows for simple import of data into statistical analysis packages. Additional studies are recommended to identify the constructs underlying student interaction, standardize the instrument, and to examine student expectations of the benefit of distance education.

ADDRESS FOR CORRESPONDENCE

The ICCE instrument may be used without cost. Please contact John Perrine, MD at japerrine@comcast.net or Judith Logan, MD at loganju@ohsu.edu.

REFERENCES

Abdal-Haqq I, Constructivism in teacher education: considerations fro those who would link practice and theory, *ERIC Digest*, ERIC-RIEO, 1999

Aday L. *Designing and Conducting Health Surveys.* Jossey-Bass, San Francisco, CA 1996

American Distance Education Consortium. *ADEC Guiding Principles for Distance Teaching and Learning* [online] http:// www.adec/admin/papers/distance-teaching_principles.html, 2002

Anderson LP, Banks SR and Leary PA. The effect of interactive television on student satisfaction, *Journal of Education for Business*, Jan./Feb. 2002

Berner ES and Adams B. Added value of video compared to audio lectures for distance learning, *Proceedings of the International Medical Informatics Association Education Conference*, April, 2003

Bonk C and Reynolds T Learner-centered Web instruction for higher-order thinking, teamwork and apprenticeship. In B. H. Kahn (Ed.), *Web-based instruction.* Englewood Cliffs: Educational Technology Publications. 1997

Brown G. *Flashlight at Washington State University* [online] http://www.ctlt.wsu.edu/resources/publications/flcases.htm 1998

Brown JS and Duguid, P. *The Social Life of Information*, Boston: Harvard Business School Press. 2000

Campus Computing Survey [online] http://www.campuscomputing.net/2003. 2002

Chickering L and Ehrmann S. Seven Principles of Good Practice in Distance Education [online] http://www.aahe.org/technology/ehrmann.htm, 1996

Cox J. *How to build the best questionnaires in the field of education,* Corwin Press, Thousand Oaks, CA, 2002

Derry JO, *et al. Purdue instructor and course appraisal: The cafeteria system.* Lafayette, IN; Purdue Research Foundation, 1974

Diaz D and Cartnal R. Students' learning styles in two classes: online distance learning and equivalent on-campus, Educational Technology, *Research and Development*, 45(5), 1997

Doll WJ and Torkzadeh T. The measurement of end-user computing satisfaction, *MIS Quarterly*, June, 1988

Ehrmann S and Zuniga R. *The Flashlight Evaluation Handbook, Version 1.0*, Washington, DC: Corporation for Public Broadcasting publication, 1997

Forelle C. Elite colleges finally embrace online degree courses, *Wall Street Journal,* Jan. 15, 2003

Gagne M and Shepherd M. Distance learning in accounting: a comparison between a distance and traditional graduate accounting class, *T.H.E. Journal* April, 28(9), 2001

Hillman DC, *et al.* Learner-interface interaction in distance education: An extension of contemporary models and strategies for practitioners. *The American Journal of Distance Education, 8*(2), 1994

Hiltz, SR. *Measuring the Importance of Collaborative Learning for the Effectiveness of ALN: A Multi-Measure, Multi-Method Approach* [online] http://www.aln.org/alnweb/journal/Vol4issue2/le/hiltz/le-hiltz.htm, 2002

Hirumi A. A framework for analyzing, designing and sequencing planned eLearning interactions, *The Quarterly Review of Distance Education*, 3(2), 2002

Huang HM, Toward constructivism for adult learners in online learning environments, *British Journal of Education Technology*, 33:1, 2002

Johnson SD, Aragon SR, Shaik N and Palma-Rivas N. Comparative analysis of learner satisfaction and learning outcomes in online and face-to-face learning environments, *Journal of Interactive Learning Research*, 11(1), 2000.

Kerlinger, FN. *Behavioral Research: A Conceptual Approach.* Holt, Rinehart and Winston, NY, 1978

Kortemeyer G and Bauer W. Multimedia collaboration content creation: The MSU Lecture Online System, *Journal of engineering Education*, Oct. 1999

Laurillard D and Taylor J. Designing the stepping stones: An evaluation of interactive media in the classroom, *Journal of Educational Television*, 20(3), 1994

Levin SR, Levin JA, Buell JG and Waddoups GL. Curriculum, Technology, and Education Reform (CTER) online: Evaluation on an online Master of Education program. *Tech Trends*. 46 (3:30-8) 5/6 2002

McIsaac MS, *et al.* Student and Teacher Perceptions of interaction in online computer-mediated communication, *Education Media International*, 36(2) 1999

Michigan Virtual University. *The Michigan Virtual University Online Course Evaluation Tool* [online] http://standards.mivu.org, 2003

Moore MG. Editorial: Three types of interaction. *The American Journal of Distance Education, 3*(2), 1989

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

OHSU, *Department of Medical Informatics & Clinical Epidemiology Master of Science Program,* [online] available at: http://www.ohsu.edu/dmice/ms/, 2003

Open University, The *OU Student Experience*, [online] available at: http://www.open.ac.uk/experience/index.htm, 2003

Oxford R, Constructivism: shape-shifting, substance and teacher education applications, *Peabody Journal of Education*, 72:1, 1997.

Palloff R and Pratt K. *Building Learning Communities in Cyberspace*, San Francisco, Jossey-Bass. 1999

Parry S and Dunn L. Benchmarking as a meaning approach to learning in online settings, *Studies in Continuing Education*, 22(2) 2000

Piaget J. *The Construction of Reality in the Child*, Basic Books, 1954

Richardson V. Constructivist teaching and teacher education: theory and practice, *Constructivist Teacher Education: Building New Understandings*, Falmer Press, 1997

Rumble G. Re-inventing distance education, 1971-2001, *International Journal of Lifelong Education*, 20:1/2 Jan.-April 2001

Russell TL. *The "No Significant Difference Phenomenon"* [online] available at http://teleeducation.nb.ca/nosignificantdifference 2003

Stanford University, *BioInformatics Training Programs*, [online] available at: http://smi.stanford.edu/academics/degrees.html, 2003

University of Phoenix, *University for Working Adults,* [online] available at: http://www.phoenix.edu, 2003

Vygotsky L. *Mind in Society* (Cole, *et al.* 9th ed.) Cambridge, MA: Harvard University Press, 1978

Zuniga RE. *The Flashlight Program for the Study and Improvement of*

Teaching and Learning with Technology [online] available at
http://www.tltgroup.org/resources/FL_and_Mt_Royal.html 2003

APPENDICES

Appendix 1. Observations and recommendations from the distance education literature, testable qualities, and questions, categorized by topic (red text)

Observation	Interaction quality	Question to test interaction quality	Reference	Cited Reference
ASSESSMENT				
If grading is not appropriately allocated to online interaction, students won't bother.	Grading should reflect the amount of time and effort put into group activities.	Grading put adequate weight on group activities (discussion groups and team projects).	Parry and Dunn (2000)	Naidu (1997)
Self-regulating one's learning is important.	Students should be able to monitor their progress throughout the course.	I knew how well I was doing throughout the class.	Northrup (2002)	
Spacing quizzes throughout the semester improves overall understanding and retention of terminology and concepts.	Graded activities should occur throughout the term.	Graded activities (e.g. tests, papers) were well spaced throughout the course.	Northrup (2002)	
Evaluations should take into account quantity of posts, the quality of participation in online discussion, course assignments and class exercises.	As stated.	Grading for participation in discussion groups was appropriate for the amount of effort.	Northrup (2002)	
STUDENT-STUDENT INTERACTION				
Collaboration and conversation is important.	As stated.	I got to know other students in the class.	Northrup (2002)	
From the online learners point of view, too much interaction may be perceived as busywork and lead to frustration.	Student-student interaction should not be overly time consuming.	Interaction with other students required too much time.	Northrup (2002)	Berge (1999)
Student-student interactions analyze and interpret data, solve problems, and share information, opinions and insights.	Students should share information and resources with each other.	I shared information and resources with other students.	Hirumi (2002)	Moore (1989)
Collaboration is a satisfying characteristic of eLearning.	Collaboration should be satisfying.	I enjoyed interacting with other students in the class.	Parry and Dunn (2000)	
Collaboration, shared goals and teamwork are powerful forces in the learning process. Group activities, simulation and the use of open-ended questions … achieve these goals.	Collaboration should promote learning.	I learned a lot by working with other students in the class.	Palloff and Pratt (1999)	
Students learn from each other. Research has shown that students in distance classes take on the role of teacher more often than do students in traditional classrooms.	Collaboration should promote learning.	I learned a lot from other students in this class.	Palloff and Pratt (1999)	

Observation	Interaction quality	Question	Reference	Cited Reference
Expression of support and encouragement exchanged between students, as well as willingness to critically evaluate the work of others.	Students should exchange ideas and resources.	I got some great ideas and resources from other students.	Palloff and Pratt (1999)	
Considerations for effective student-student interaction in eLearning environments are group size, group goals, individual roles and responsibilities, group and individual accountability, contact information, communications and grading.	Team/discussion groups should not be too big.	The size of my team or discussion group was just about right.	Hirumi (2002)	Moore (1989)
Need for social connection almost supersedes the content-oriented goals of the course. Students need to gather in cyberspace, just as they do on campus. To do this they need to establish a sense of presence online that allows their personality to come through.	Social interaction among students should be encouraged.	I got to know some of my classmates pretty well.	Palloff and Pratt (1999)	Nipper (1989)
In order for participants to connect with each other, there must be a sense of safety and trust.	Trust-building should be part of collaborative learning.	I had a bond of trust with some of my classmates.	Palloff and Pratt (1999) ch. 11	
COURSE DESIGN				
Ideal eLearning engages community members in deep thinking, provides multiple viewpoints, supports reflection and offers frequent feedback and guidance toward higher standards.	Course work should promote deep thinking and multiple viewpoints.	The course got me interested in exploring new ideas and viewpoints.	Tu (2002)	
Good online teaching is student-centered, self-reflective, and includes discussion of ideas, concepts and theories.	As stated.	I was encouraged to think and talk about ideas and theories.	Parry and Dunn (2000)	Berge (1997)
We need to think more creatively about how to develop course designs that respond to a greater variety of learning styles rather than concluding that online learning is more suitable for one type of student than another.	Different learning styles should be supported.	I was able to learn in the way that worked best for me.	Twigg (2002)	

Observation	Interaction quality	Question	Reference	Cited Reference
Students are more comfortable with traditional classroom format (lecture, reading, discussion) but favorably rate innovative strategies such as debate, role-playing and gaming.	Students should be comfortable with course format.	I was comfortable with the structure of the class.	Northrup (2002)	
Distance learning should balance the goals of user-friendliness, low user overhead and learning enhancements.	The course Web site should be easy to use.	The course Web site was easy to use.	Okamoto (2001)	
Creating structure is important to prevent students from becoming overwhelmed, or falling behind.	Course material should be well organized and presented.	The course material was well organized and presented.	Christianson et al. (2002)	
FEEDBACK				
Students should be guided and encouraged in their ability to give each other meaningful feedback on their work.	As stated.	Other students gave me constructive feedback on my work.	Palloff and Pratt (1999) ch. 6	
Feedback provides learning guidance, lesson sequence advisement, motivational messages, critical comparisons and information about answer correctness and timeliness.	As stated.	Feedback from the instructor or TA was helpful.	Hirumi (2002)	Hoska (1993)
Encouragement, praise and	Instructors should go	The instructor let me	Twigg (2002)	
Students report information overload, communication anxiety in relation to the delayed response in an asynchronous environment.	Feedback should be prompt.	Feedback from the instructor or TA was always prompt.	Palloff and Pratt (1999)	
Feedback is important: needs to at least two times per week.	There should be enou feedback.	I got enough feedback from the instructor or T/	Northrup (2002)	
STUDENT-INSTRUCTOR INTERACTION				
As a facilitator, the instructor provides gentle guidance and a loosely constructed framework as a container for the course, thus allowing students to explore the course material, as well as related materials, without restriction. This is not done through the traditional use of lectures followed by some for form of discussion. Instead, the instructor may provide general topics within the body of knowledge.	The instructor needs to keep discussions on the right track.	The instructor kept group discussions on the right track.	Palloff and Pratt (1999) ch. 6	

Observation	Interaction quality	Question	Reference	Cited Reference
It is important to contact the students frequently to check if they are having any problems with the course, assignments, use of technology, and eventually get their feedback for improving the course.	As stated.	The instructor checked in to make sure I was doing okay. - OR - The support staff checked in to make sure I was doing okay.	McIsaac (2000)	
Instructors assume a dual role, assist students in the accomplishment of learning objectives and flexible facilitation. There should be an effort to address the individual needs of students, but within an established framework.	As stated.	The instructor was very helpful. - OR - The support staff were very helpful.	Christianson et al. (2002)	
Activities recommended for moderation: participate in discussion regularly, express honest opinions, engage peer moderations; allow reasonable venting; show concern and support for community.	The instructor should be a regular participant in group discussions.	The instructor was a regular participant in group discussions.	Tu (2002)	
MEDIA				
It is necessary to seek the media mix that will increase the effectiveness of the learning process.	As stated.	The mix of media in the classes (audio, text or video) provided an effective learning environment.	Okamoto (2001)	
Totally self-paced models (of distance education) do not work well and can lead to high attrition rates.	There should be a balance between free and structured learning.	There was a good balance of structured learning and free exploration.	Twigg (2002)	
As interactive technologies become the staple communication vehicle for innovative virtual worlds, effective GUI design will ensure that learner focus is on learning rather than operating the software.	Learning materials should be easy to use.	Learning materials were easy to use.	Metros (2002)	
Poor interface design may generate high cognitive demands which take attention away from learning	The interface should be user-friendly.	The course Web site was user-friendly.	Hirumi (2002)	Metros and Hedberg (2002)
When information is complex, has high element interactivity, or when different sources need to be integrated, graphic-auditory presentation is better than graphic-textual presentation.	Course lectures should be able to convey complex information. To make complex information easier to understand, it should be presented in a richer (audio-video) media format.	Course lectures were in a suitable format (audio, text or video) to convey complex information.	Tuovinen (2000)	Mousavi, et al. (1995)

Observation	Interaction quality	Question	Reference	Cited Reference
Multimedia presentation is better when material is complex. If the information is simple, text is better.		I liked the way the lectures were presented (audio, text or video).	Tuovinen (2000)	
Multimedia-oriented learning can be more user-friendly.	Course lectures should be user-friendly.	Course lectures were easy to understand.	Okamoto (2001)	
Creating appropriate challenges that are effectively represented by the visual design requires understanding of how learners need to access and manipulate available resources and an appreciation for the skills and knowledge of each design team member.	Course material should be engaging.	The course materials were engaging.	Metros (2002)	
The alternative forms of representation available in video can be particularly important for conceptual understanding that requires students to think in a particular way about an idea.	Course material should be easy to understand.	The way course materials were presented (audio, text or video) made them easy to understand.	Laurillard and Taylor (1994)	
Increasing numbers on online courses ask learners to review external websites, as well as communicate with others outside of class to promote knowledge construction and social discourse. Such interactions include exchanges with TAs, mentors and subject matter experts.	Students should be encouraged to seek outside sources of information (e.g. Web sites, students and experts at other institutions).	I was encouraged to explore outside sources (e.g. Web sites, students and experts at other institutions).	Hirumi (2002)	Bonk and King (1998)

TECHNOLOGY

Students are frustrated when technology does not perform adequately.	The course should be free of technical problems.	I had problems with course technology.	Northrup (2002)	
Comfort with technology contributes to greater sense of well-being and likelihood of participation.	Students should be comfortable with course technology.	I was comfortable with the technology.	Palloff and Pratt (1999)	
Technology should be simple to operate.	As stated.	Course technology was user-friendly.	Palloff and Pratt (1999)	

Observation	Interaction quality	Question	Reference	Cited Reference
Technology should be user-friendly, visually appealing, easy to navigate.	As stated.	The course Web site was easy to navigate.	Palloff and Pratt (1999)	
Designers and instructors should provide clear expectations for online collaboration.	As stated.	The procedures for working with other students were clear.	Northrup (2002)	
Collaborative learning should be explicitly organized.	As stated.	Team projects and discussion groups were well planned and organized.		Bourdeau and Bates (1997)
Clear and specific details about the course structure, assignments, activities and evaluation are needed, along with concise instructions for navigating the online environment.	As stated.	Expectations about assignments and activities were clear.	Berge (2002)	Lansdell (2001)
Need structured times that assignments are due.	As stated.	I knew what time activities and assignment were due.	Northrup (2002)	
GOALS AND OBJECTIVES				
The expectations for the course should be made clear, both in terms of learning activates and competencies to be acquired. This goes beyond providing a syllabus.	As stated.	Assignments and course requirements were clear.	Berge (2002)	
It is important to package information for students in an organized structure.	As stated.	Course materials were well organized.	Christianson et al. (2002)	Halstead and Coudret (2000)
Encourage self-awareness of the knowledge construction process.	Students should learn something about the learning process.	I learned about the process of learning.	McLoughlin (2002)	
MULTIPLICITY				
The more demanding and complex the content, the more beneficial the multimodal interaction in content presentation, instructor-student interaction and learner-learner interaction.	There should be diversity in learning materials and activities.	The course had a good variety of learning materials and activities.	Tuovinen (2000)	
Provide experience in and appreciation of multiple perspectives	As stated.	I learned from a good variety of perspectives.	McLoughlin (2002)	

Observation	Interaction quality	Question	Reference	Cited Reference
New learning environments are characterized by flexible enrollment options; personalized, on-demand, 24/7 student services; innovative curricular design that includes a focus on applied or problem-based learning taught by practicing professionals; and learner assessment that is integrated throughout the curriculum by diagnosing the students' knowledge and skill levels as they begin their programs of study and by responding accordingly.	Students should be encouraged to solve real-world problems.	In this class I was able to participate in solving relevant, "real-world" problems.	Twigg (2002)	
The essence of inquiry is when the student is personally challenged with a problem to solve, a project to complete, or a dilemma or resolve.	As stated.	I felt challenged and motivated by this class.	Berge (2002)	
RELEVANT CONTENT				
Placing learning in a meaningful context is critical.	As stated.	The subject matter was personally relevant.	Berge (2002)	
The types of questions asked to kick off the discussion of a topic within a course can encourage students to bring their life experiences into the classroom.	Discussions should include students' personal experiences.	I was able to relate personal experiences in class discussions.	Palloff and Pratt (1999) ch. 8	
Learning should be become more demand-driven. People learn in response to need. When students cannot see the need for what is being taught, they ignore or fail to assimilate.	Students should see tangible benefit from what they are learning.	I saw real uses for what I was learning.	Cogburn, et al. (2001)	Brown and Duguid (2000)
SUPPORT				
Timeliness of response is a major indicator of support.	Instructor support should be prompt.	The instructor answered my questions right away.	Northrup (2002)	
Having a mentor in place to provide assistance is important.	The instructor should be accessible.	The instructor was accessible.	Northrup (2002)	
Learners need access to support and service from staff such as librarians, advisors and counselors.	As stated.	I had good access to library resources.	Hirumi (2002)	
Quick assess to technical support staff is essential.	As stated.	Technical support was accessible and quick to respond.	Hirumi (2002)	

Observation	Interaction quality	Question	Reference	Cited Reference
Access to a large support system of fellow students and tutors who are available virtually around the clock is a key component to these new designs.	As stated.	Other students were able to answer many of my questions.	Twigg (2002)	
Social support, peer support and task support represent core elements for the learning process in distance education.	As stated.	Other students were very supportive.	McLoughlin and Oliver (1998)	
STUDENT-CONTENT INTERACTION				
Recent advances in communications technologies and in pedagogy envisage an active, participatory role for students, an as initiators and co-participants in self-regulating learning processes.	Students should be active participants in learning.	I actively participated, rather than just learned information.	Collins and Moonen (2001)	
Students need to interact with learning materials that allow them greater choices of assignments and resources. The key goal is for the students to become engaged in active "doing" in the learning process -- that is, to move beyond merely reading test.	As stated.	This course allowed me a choice of interactive learning materials (e.g. Web sites, audio/visual materials).	Twigg (2002)	
As the control of learning shifts from teacher to learner, and as the value of the student's time becomes more important, individualized learning becomes critical.	Students should be able to individualize their learning.	I was able to pick the learning materials which worked best for me.	Berge (2002)	
Question to ask student regarding online activities: The activity was an enjoyable way to learn.	Students should enjoy interacting with course materials.	I enjoyed interacting with class learning materials (Web sites, audio-visual material, discussion boards).	Lawless and Freake (2001)	
CLASS DISCUSSION				
Participants reported liking to discuss ideas, concepts and information with peers.	As stated.	I enjoyed discussing class ideas with other students.	Northrup (2002)	
Members can become confused and overloaded if guidelines for participation are not established at the start.	Guidelines for discussion groups should be clear.	Guidelines for discussion were clear.	Palloff and Pratt (1999)	

Observation	Interaction quality	Question	Reference	Cited Reference
The instructor needs to monitor discussions and jump in as facilitator if necessary (e.g. for unequal participation).	As stated.	The instructor kept discussions on the right track.	Palloff and Pratt (1999)	
The introverted student who may not feel comfortable speaking out or asking for help in a face-to-face setting my flourish online.	Students should feel comfortable in discussion groups	I felt comfortable participating in discussion groups.	Palloff and Pratt (1999) ch. 11	
Instructor participation in discussion groups (moderated by students) adds more credibility to the discussion. Students want to hear from instructors.	The instructor should be an active participant in class discussions.	The instructor added valuable insights to the discussions.	McIsaac (2000)	

Appendix 2. Modified questions for content and distance education experts, categorized by Moore-Hirumi interaction categories

Item Code	Item Description	Response Category
	Student-Classmate interaction:	
S1	I enjoyed interacting with classmates.	Likert
S2	I got to know classmates personally.	Likert
S3	I gained a level of trust with classmates.	Likert
S4	I shared resources with classmates.	Likert
S5	Classmates supported me with my class work.	Likert
S6	I worked with fellow students on class assignments.	Likert
S7	Classmates gave me valuable feedback on my work.	Categorical (ordinal)
S8	I learned through participating in class discussions.	Likert
S9	Students in the class motivated me.	Likert
S10	I was an active participant in class discussions.	Likert
S11	I was comfortable working with my project group.	Likert
S12	In my group, there was an equitable distribution of work.	Likert
S13	There was adequate weight put on the grading of my group project.	Likert
S14	I had contact with classmates: (More than 10 times during the term; Six to 10 times during the term; Two to 5 times during the term; Once during the term; I didn't have contact with classmates).	Categorical (ordinal)
S15	My interaction with classmates occurred mostly through: (Email; The class discussion board; Phone; A variety of media.)	Categorical
	Student-Instructor interaction:	
T1	The instructor was encouraging.	Likert
T2	The instructor valued my class contributions.	Likert
T3	The instructor added valuable insights to class discussions	Likert
T4	The instructor kept class discussions on the right track.	Likert
T5	I received adequate feedback from the instructor.	Likert
T6	The instructor answered my questions in a timely fashion.	Likert
T7	The instructor was accessible.	Likert
T8	I enjoyed my interactions with the course instructor.	Likert
T9	My contact with the instructor was: (More than five times during the term; Two to four times during the term; Once during the term; I didn't have contact with the instructor.)	Categorical (ordinal)
T10	My interaction with the instructor occurred mostly through: (Email; The class discussion board; Phone; A variety of media.)	Categorical
	Student-Course Material Interaction:	
C1	Lectures were relevant to the course.	Likert
C2	Reading material was relevant to the course.	Likert
C3	The course Web site was easy to use.	Likert
C4	Course lectures were easy to access/view.	Likert

C5	Course lectures were easy to understand.	Likert
C6	The course discussion board was easy to use.	Likert
C7	I enjoyed watching the course lectures.	Likert
C8	I enjoyed participating in course discussions.	Likert
C9	I enjoyed the assigned reading material.	Likert
C10	I learned a lot by watching the course lectures.	Likert
C11	I learned a lot through course readings.	Likert
C12	My preference for lecture format is: (Video with synchronized PowerPoint slides; Audio with synchronized PowerPoint slides; Video with hard-copy notes; Audio with hard-copy notes; I generally did not view the lectures.)	Categorical
	Student-Instruction Interaction:	
N1	I had contact with course technical support: (More than five times during the term; Two to four times during the term; Once during the term; I didn't have contact with technical support.)"	Categorical (ordinal)
N2	I am satisfied with the technical support I received.	Likert
N3	The technical support staff answered my questions in a timely fashion.	Likert
N4	The technical support staff was accessible.	Likert
N5	I learned a lot through taking the quizzes/exams.	Likert
N6	I knew how well I was doing throughout the class.	Likert
N7	I was able to relate personal experiences/knowledge to class discussions.	Likert
N8	The subject matter was relevant to me.	Likert
N9	I saw real uses for what I was learning.	Likert
N10	The mix of media in the classes made for a good learning environment.	Likert
N11	I was able to choose learning materials that worked well for me.	Likert
N12	Class discussions expanded upon the ideas and theories presented in class.	Likert
N13	I felt challenged by this class.	Likert

Appendix 3A. Item-item correlation matrix for student-classmate interactions

Item Code	Item Description	SG1	SG2	S1	S2	S3	DX	S5	S6	S7	SX
SG1	I am satisfied with my interactions with students in this class.	1.00									
SG2	I benefited from interactions with students in the class.	0.66	1.00								
S1	I enjoyed interacting with classmates.	0.57	0.81	1.00							
S2	I got to know classmates personally.	0.34	0.29	0.25	1.00						
S3	I worked with fellow students on class assignments.	-0.08	0.01	0.01	-0.22	1.00					
DX	I participated in class discussions.	0.18	0.22	0.10	-0.02	-0.04	1.00				
S5	Classmates gave me feedback on my work.	0.40	0.40	0.44	0.21	-0.20	0.23	1.00			
S6	I learned through participating in class discussions	0.52	0.75	0.69	0.00	0.06	0.23	0.35	1.00		
S7	I enjoyed participating in class discussions.	0.51	0.67	0.73	0.10	0.09	0.19	0.30	0.79	1.00	
SX	I interacted with classmates.	0.17	0.15	0.07	-0.11	-0.03	0.32	0.11	0.12	0.12	1.00

Green text signifies correlation ≥ .40 significant at $p < 0.05$ (two-tailed); red text signifies a correlation without statistical significance

Appendix 3B. Item-item correlation matrix for student-instructor interactions

Item Code	Item Description	TG1)	TG2	T1	T2	T3	T4	T5	TX
TG1	I am satisfied with my interactions with the instructor.	1.00							
TG2	I benefited from my interactions with the instructor.	0.78	1.00						
T1	The instructor valued my class contributions.	0.61	0.68	1.00					
T2	The instructor added valuable insights to class discussions.	0.58	0.55	0.45	1.00				
T3	The instructor kept class discussions on the right track.	0.54	0.66	0.53	0.82	1.00			
T4	The instructor answered my questions in a timely fashion.	0.75	0.68	0.47	0.64	0.62	1.00		
T5	The instructor was accessible.	0.77	0.66	0.50	0.71	0.65	0.89	1.00	
TX	I interacted with the instructor.	0.28	0.47	0.29	0.11	0.26	0.16	0.03	1.00

Green text signifies correlation ≥ .40 significant at $p < 0.05$ (two-tailed); red text signifies a correlation without statistical significance

Appendix 3C. Item-item correlation matrix for student-learning material interactions

Item Code	Item Description	MG1	MG2	ML1	ML2	ML3	LX	ML5	MR7	ML8	MR9
MG1	I'm satisfied with course learning materials.	1.00									
MG2	I learned through interacting with course materials.	0.75	1.00								
ML1	Lectures were relevant to course objectives.	0.63	0.47	1.00							
ML2	Course lectures were easy to access/view.	0.45	0.44	0.35	1.00						
ML3	Course lectures were easy to understand.	0.52	0.53	0.61	0.59	1.00					
LX	I viewed the course lectures.	0.34	0.33	0.31	0.09	0.27	1.00				
ML5	I enjoyed watching the course lectures.	0.51	0.62	0.50	0.54	0.63	0.36	1.00			
MR7	I enjoyed the assigned reading material.	0.73	0.56	0.60	0.37	0.41	0.32	0.52	1.00		
ML8	I learned a lot by watching the course lectures.	0.55	0.65	0.66	0.51	0.68	0.33	0.80	0.50	1.00	
MR9	I learned by reading assigned materials.	0.78	0.68	0.62	0.50	0.62	0.34	0.48	0.75	0.58	1.00

Green text signifies correlation ≥ .40 significant at $p < 0.05$ (two-tailed); red text signifies a correlation without statistical significance

Appendix 3D. Item-item correlation matrix for student-instruction interactions

Item Code	Item Description	CG1	CG2	CK1	CK2	CK3	CS4	CS5	CS6	CS7
CG1	I'm satisfied with the class, overall.	1.00								
CG2	The course was successful.	0.89	1.00							
CK1	I interacted with technical support.	0.43	0.35	1.00						
CK2	I am satisfied with the technical support I received.	0.42	0.32	0.92	1.00					
CK3	The technical support staffed answered my questions In a timely fashion.	0.40	0.31	0.91	1.00	1.00				
CS4	I knew how well I was doing throughout the course.	0.54	0.46	0.23	0.25	0.23	1.00			
CS5	I saw real uses for what I was learning.	0.73	0.81	0.32	0.24	0.24	0.49	1.00		
CS6	I was able to choose learning materials that worked well for me.	0.65	0.65	0.29	0.33	0.32	0.50	0.56	1.00	
CS7	I felt challenged by the course.	0.63	0.67	0.28	0.31	0.30	037	0.50	0.52	1.00

Green text signifies correlation ≥ .40 significant at $p < 0.05$ (two-tailed); red text signifies a correlation without statistical significance

Appendix 4. Corrected item-total correlations

Item Code	Item Description	Corrected Item-Total Correlation	Item Code	Item Description	Corrected Item-Total Correlation
	Student-classmate interaction			**Student-instructor interaction**	
SG1	I am satisfied with my interactions with students in this class.	.67	TG1	I am satisfied with my interactions with the instructor.	.80
SG2	I benefited from interactions with students in the class.	.84	TG2	I benefited from my interactions with the instructor.	.82
S1	I enjoyed interacting with classmates.	.79	T1	The instructor valued my class contributions.	.64
S2	I got to know classmates personally.	.19	T2	The instructor added valuable insights to class discussions	.74
S3	I worked with fellow students on class assignments.	-.06	T3	The instructor kept class discussions on the right track.	.76
DX	I participated in class discussions.	.25	T4	The instructor answered my questions in a timely fashion.	.80
S5	Classmates gave me feedback on my work.	.45	T5	The instructor was accessible.	.82
S6	I learned through participating in class discussions	.73	TX	I interacted with the instructor.	.27
S7	I enjoyed participating in class discussions.	.71			
SX	I interacted with classmates.	.14			
	Student-learning material interaction			**Student-instruction interaction**	
MG1	I'm satisfied with course learning materials.	.77	CG1	I'm satisfied with the class, overall.	.74
MG2	I learned through interacting with course materials.	.73	CG2	The course was successful.	.68
ML1	Lectures were relevant to course objectives.	.69	CT1	I interacted with technical support.	.76
ML2	Course lectures were easy to access/view.	.54	CT2	I am satisfied with the technical support I received.	.72
ML3	Course lectures were easy to understand.	.69	CT3	The technical support staffed answered my questions in a timely fashion.	.71
LX	I viewed the course lectures.	.40	CS4	I Knew how well I was doing throughout the course.	.46
ML5	I enjoyed watching the course lectures.	.72	CS5	I saw real uses for what I was learning.	.58
MR7	I enjoyed the assigned reading material.	.66	CS6	I was able to choose learning materials that worked well for me.	.60
ML8	I learned a lot by watching the course lectures.	.76	CS7	I felt challenged by the course.	.55
MR9	I learned by reading assigned materials.	.77			

Green text signifies correlation ≥ .40 significant at $p < 0.05$ (two-tailed); red text signifies a correlation without statistical significance

Appendix 5A. Items Removed from Interaction-Centered Evaluation Tool, Correlation test Scores

Item Code	Item Description	Test Statistic	Corrected Item-Total Correlation	Item-Global Item (CG1) Correlation	Item-Global Item (CG2) Correlation
	Student-classmate				
S2	I got to know classmates personally.	Pearson r correlation	.19	0.34	0.29
S3	I worked with fellow students on class assignments.	Pearson r correlation	-.06	-0.08	0.01
DX	I participated in class discussions.	Pearson r correlation	.25	0.18	0.22
SX	I interacted with classmates.	Pearson r correlation	.14	0.17	0.15
	Student-instructor				
TX	I interacted with the instructor.	Pearson r correlation	.27	0.28	0.47
	Student-learning material				
LX	I viewed the course lectures.	Pearson r correlation	.40	0.34	0.33
ML10	My preference for the lecture format is (or would have been): Audio with PowerPoint slides, Audio with course handout, etc.	Not a scale question	N/A	N/A	N/A
	Student-instruction				
CK1	I interacted with technical support.	Pearson r correlation	.76	0.43	0.35
CK2	I am satisfied with the technical support I received.	Pearson r correlation	.72	0.42	0.32
CK3	The technical support staffed answered my questions	Pearson r correlation	.71	0.40	0.31

Green text signifies correlation ≥ .40 significant at $p < 0.05$ (two-tailed); red text signifies a correlation without statistical significance.

Appendix 5B. Items Removed from Interaction-Centered Evaluation Tool, Median test Scores by grouping variable and scoring item.

Item Code	Item Description (Grouping Variable)	Test Type, (Statistic)	SG1	SG2	TG1	TG2	MG1	MG2	CG1	CG2
	Demographic									
DE1	How many distance courses have you taken previously?	Chi Square, Significance	2.49 .478	4.63 .201	3.07 .381	1.73 .631	3.34 .342	4.64 .200	.434 .933	.728 .867
DE2	What is your academic background?	Chi Square, Significance	3.64 .457	2.60 .626	5.45 .142	4.98 .173	.997 .910	3.29 .510	3.42 .490	3.36 .499
DE3	What is your computing experience?	Chi Square, Significance	2.41 .121	.092 .762	6.11 .013	1.49 .229	1.33 .249	.036 .849	.028 .867	.296 .586
	Type-of-interaction									
SMem	My interaction with the other students occurred through email.	Chi Square, Significance	2.10 .147	.818 .366	.255 .614	1.37 .242	2.27 .131	.112 .738	.377 .539	.007 .933
SMcd	My interaction with the other students occurred through class discussion boards.	Chi Square, Significance	N/A	N/A	N/A	N/A	.163 .687	.461 .497	.353 .552	.406 .524
SMph	My interaction with the other students occurred through phone.	Chi Square, Significance	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SMpe	My interaction with the other students occurred through personal contact.	Chi Square, Significance	.908 .341	.173 .677	.003 .960	2.58 .108	.379 .583	.203 .652	.823 .364	.510 .475
TMem	My interaction with the instructor occurred through email.	Chi Square, Significance	.051 .821	.508 .476	1.29 .255	1.94 .163	1.08 .297	.781 .377	2.38 .123	2.74 .098
TMcd	My interaction with the instructor occurred through class discussion boards.	Chi Square, Significance	10.76 .001	2.43 .119	.147 .702	.677 .411	1.15 .284	3.32 .068	2.52 .112	2.91 .088
TMph	My interaction with the instructor occurred through phone.	Chi Square, Significance	1.85 .174	.446 .504	.440 .507	.657 .418	.163 .687	.461 .497	.353 .552	.406 .524
TMpe	My interaction with the instructor occurred through personal contact.	Chi Square, Significance	4.19 .042	.157 .692	.182 .670	2.59 .107	3.57 .059	.118 .731	3.06 .080	2.26 .132

Top score represents Chi-Square value; bottom represents significance. Green text signifies a Chi Square value significant at $p < 0.05$ (two-tailed); red text signifies a Chi Square value without statistical significance. N/A signifies not enough valid cases to perform Median Test (e.g. few students reported corresponding with other students by phone).

Appendix 6. Test-retest reliability (Spearman's *rho*)

Item Code	Item Description	Test/Retest Score (Spearman's rho)	Significance *p* (two-tailed)
	Student-instruction (overall course) interaction:		
CG1	I'm satisfied with the class, overall.	.82	.012
CG2	The course was successful.	.93	.001
	Student-student interaction:		
SG1	I am satisfied with my interactions with students in this class.	.69	.085
SG2	I benefited from interactions with students in the class.	.64	.12
	Student-instructor interaction:		
TG1	I am satisfied with my interactions with the instructor.	0.0	1.0
TG2	I benefited from my interactions with the instructor.	.61	.20
	Student-content interaction:		
MG1	I'm satisfied with course learning materials.	.80	.017
MG2	I learned through interacting with course materials.	.59	.12

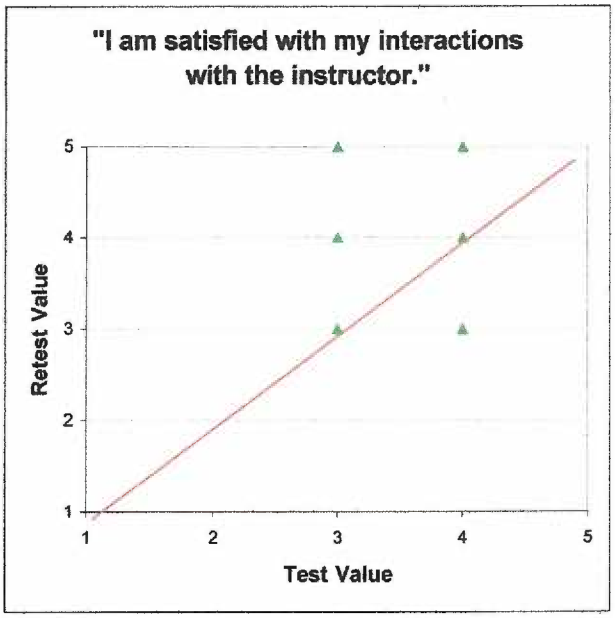
Green text signifies correlation ≥ .80 significant at $p < 0.05$ (two-tailed); red text signifies a correlation without statistical significance

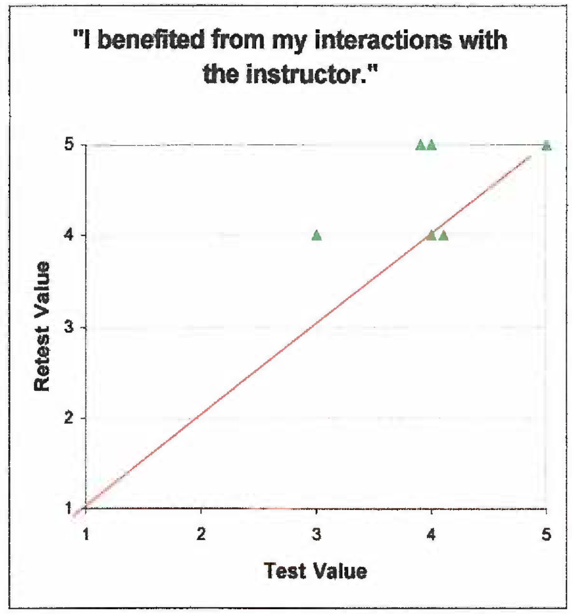
Appendix 7. ICEE item correlation with course satisfaction (CG1) and perceived course benefit (CG2)

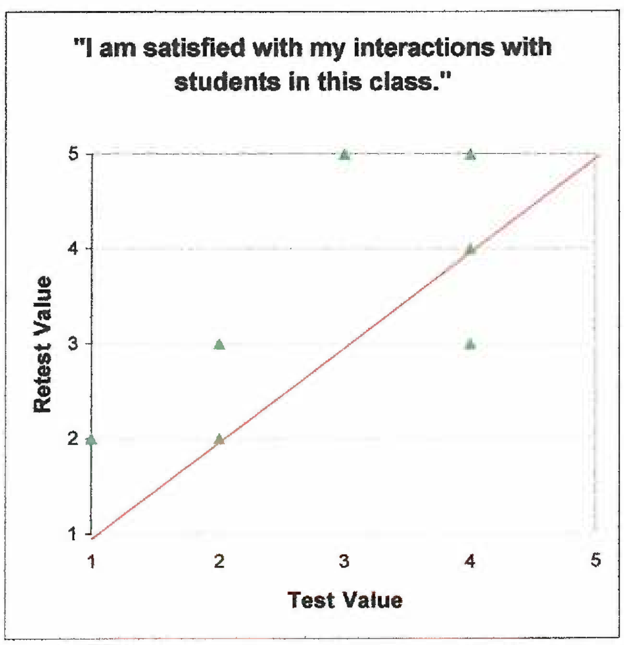
Item Code	Item Description	CG1–Satisfied (global)	CG2–Benefited (global)
	Student-student interaction:		
SG1	I am satisfied with my interactions with students in this class.	0.44	0.46
SG2	I benefited from interactions with students in the class.	0.53	0.55
S1	I enjoyed interacting with classmates.	0.43	0.50
S5	Classmates gave me feedback on my work.	0.21	0.26
S6	I learned through participating in class discussions	0.63	0.61
S7	I enjoyed participating in class discussions.	0.69	0.62
	Student-instructor interaction:		
TG1	I am satisfied with my interactions with the instructor.	0.56	0.56
TG2	I benefited from my interactions with the instructor.	0.65	0.62
T1	The instructor valued my class contributions.	0.50	0.49
T2	The instructor added valuable insights to class discussions.	0.61	0.56
T3	The instructor kept class discussions on the right track.	0.65	0.62
T4	The instructor answered my questions in a timely fashion.	0.60	0.50
T5	The instructor was accessible.	0.60	0.53
	Student-learning material interaction:		
MG1	I'm satisfied with course learning materials.	0.63	0.58
MG2	I learned through interacting with course materials.	0.67	0.70
ML1	Lectures were relevant to course objectives.	0.56	0.49
ML2	Course lectures were easy to access/view.	0.30	0.19
ML3	Course lectures were easy to understand.	0.52	0.51
ML5	I enjoyed watching the course lectures.	0.41	0.46
MR7	I enjoyed the assigned reading material.	0.50	0.49
ML8	I learned a lot by watching the course lectures.	0.48	0.51
MR9	I learned by reading assigned materials.	0.70	0.70
	Student-instruction interaction:		
CS4	I knew how well I was doing throughout the course.	0.56	0.50
CS5	I saw real uses for what I was learning.	0.70	0.78
CS6	I was able to choose learning materials that worked well for me.	0.63	0.64
CS7	I felt challenged by the course.	0.70	0.71
TINX	**Total interaction (sum of SX, DX, TX and LX)**	0.43	0.46

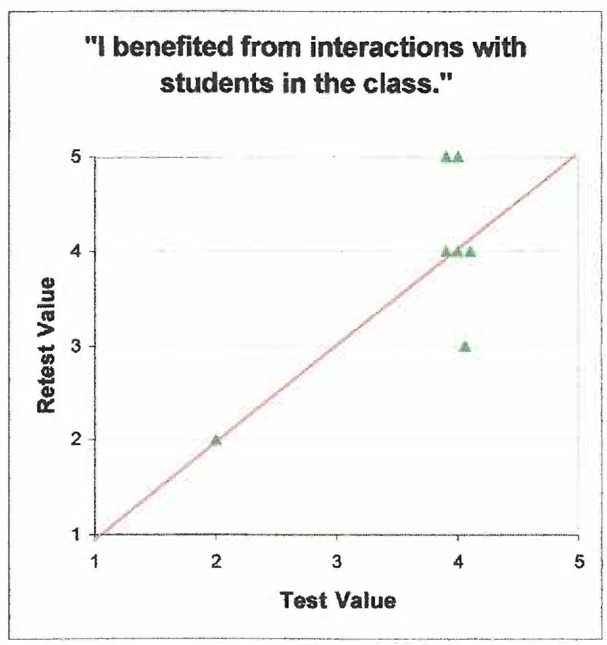
Green text signifies correlation ≥ .40 significant at $p < 0.05$ (two-tailed); red text signifies a correlation without statistical significance

Appendix 8. Test (x-axis) vs. retest (y-axis) score graphs for global items (TG1, TG2, SG1 and SG2) – diagonal lines indicate equal test-retest scores; points above the diagonal indicate retest scores greater than test scores. The pattern of retest scores ≥ test scores suggests a systematic error in the testing method.









Appendix 9. Descriptive statistics for questionnaire items by course

Item Code	Questionnaire Item	All Courses			Art of HIPAA			Biometry		
	Student-Instructor Interaction	N	Mean	S.D.	N	Mean	S.D	N	Mean	S.D.
S1	I enjoyed interacting with classmates.	51	3.84	0.83	1	3	N/A	6	4.50	0.55
S2	I got to know classmates personally.	51	2.41	0.92	1	2	N/A	6	3.17	0.98
S3	I worked with fellow students on class assignments.	51	2.86	0.35	1	3	N/A	6	2.67	0.52
DX	I participated in class discussions.	51	2.69	0.58	1	3	N/A	6	2.50	0.55
S5	Classmates gave me feedback on my work.	51	1.82	0.59	1	2	N/A	6	1.83	0.75
S6	I learned through participating in class discussions	48	3.81	0.91	1	4	N/A	6	4.17	0.75
S7	I enjoyed participating in class discussions.	48	3.60	1.12	1	3	N/A	6	4.00	0.63
SX	I interacted with classmates.	55	4.13	1.17	1	4	N/A	6	4.50	0.55
SG1	I am satisfied with my interactions with students in this class.	51	3.18	1.13	1	2	N/A	6	3.83	0.75
SG2	I benefited from interactions with students in the class.	51	3.76	0.89	1	4	N/A	6	4.50	0.55
	Student-Instructor Interaction									
T1	The instructor valued my class contributions.	46	3.57	0.78	1	3	N/A	6	3.83	0.41
T2	The instructor added valuable insights to class discussions	48	3.02	1.06	1	3	N/A	6	3.50	0.55
T3	The instructor kept class discussions on the right track.	48	3.02	1.21	1	4	N/A	6	3.83	0.98
T4	The instructor answered my questions in a timely fashion.	46	3.80	1.13	1	4	N/A	6	4.00	0.89
T5	The instructor was accessible.	46	3.89	1.04	1	4	N/A	6	4.33	0.52
TX	I interacted with the instructor.	55	2.84	0.98	1	3	N/A	6	3.33	0.52
TG1	I am satisfied with my interactions with the instructor.	46	3.57	0.98	1	3	N/A	6	4.00	0.89
TG2	I benefited from my interactions with the instructor.	46	3.70	1.03	1	4	N/A	6	4.00	0.89
	Student-Learning Material Interaction									
ML1	Lectures were relevant to course objectives.	45	4.07	0.84	1	4	N/A	6	4.67	0.52
ML2	Course lectures were easy to access/view.	45	3.82	0.89	1	4	N/A	6	4.33	0.82
ML3	Course lectures were easy to understand.	45	3.93	0.72	1	4	N/A	6	4.17	0.75
LX	I viewed the course lectures.	55	2.31	1.12	1	3	N/A	6	3.00	0.00
ML5	I enjoyed watching the course lectures.	45	3.60	0.99	1	4	N/A	6	4.17	0.75
MR7	I enjoyed the assigned reading material.	55	3.55	1.05	1	4	N/A	6	4.17	0.98
ML8	I learned a lot by watching the course lectures.	45	3.91	0.87	1	4	N/A	6	4.67	0.82
MR9	I learned by reading assigned materials.	55	3.76	1.04	1	4	N/A	6	4.50	0.84
MG1	I'm satisfied with course learning materials.	55	3.24	1.10	1	4	N/A	6	4.00	0.63
MG2	I learned through interacting with course materials.	55	3.67	1.07	1	4	N/A	6	4.50	0.84
	Student-Instruction Interaction									
CK1	I interacted with technical support.	55	1.64	0.85	1	1	N/A	6	2.00	1.10
CK2	I am satisfied with the technical support I received.	43	2.21	2.23	0	N/A	N/A	5	2.80	2.59
CK3	The technical support staffed answered my questions In a timely fashion.	43	2.16	2.20	0	N/A	N/A	5	2.80	2.59
CS4	I Knew how well I was doing throughout the course.	55	3.24	1.17	1	2	N/A	6	4.00	0.89
CS5	I saw real uses for what I was learning.	55	3.89	0.99	1	4	N/A	6	4.00	0.89
CS6	I was able to choose learning materials that worked well for me.	55	3.04	1.00	1	3	N/A	6	3.83	0.41
CS7	I felt challenged by the course.	55	3.51	1.25	1	5	N/A	6	4.67	0.52
CG1	I'm satisfied with the class, overall.	55	3.40	1.21	1	4	N/A	6	4.17	0.41
CG2	The course was successful.	55	3.56	1.13	1	3	N/A	6	4.33	0.52

Appendix 9 (continued). Descriptive statistics for questionnaire items by course

Item Code	Intermediate XLM			Computer Networks			Blogging the EMR			Extreme SNOMED			History of ICD-9		
	N	Mean	S.D.	N	Mean	S.D.	N	Mean	S.D.	N	Mean	S.D.	N	Mean	S.D.
S1	11	4.00	0.89	3	3.67	1.15	20	3.55	0.89	5	4.20	0.45	5	3.80	0.45
S2	11	2.18	0.87	3	1.67	0.58	20	2.20	0.83	5	2.60	0.89	5	3.20	0.84
S3	11	3.00	0.00	3	3.00	0.00	20	2.85	0.37	5	3.00	0.00	5	2.60	0.55
DX	11	2.91	0.30	3	2.00	1.00	20	2.70	0.57	5	3.00	0.00	5	2.40	0.89
S5	11	1.82	0.40	3	1.33	0.58	20	1.80	0.62	5	2.00	0.71	5	2.00	0.71
S6	11	3.91	0.54	2	4.00	1.41	19	3.47	1.22	5	4.20	0.45	4	4.00	0.00
S7	11	3.82	1.25	2	4.00	1.41	19	3.16	1.30	5	4.00	0.71	4	4.00	0.00
SX	12	4.33	1.23	4	2.50	1.29	22	4.05	1.25	5	4.40	0.89	5	4.60	0.55
SG1	11	3.27	1.01	3	2.33	0.58	20	2.90	1.21	5	3.40	1.34	5	3.80	1.10
SG2	11	3.91	0.70	3	3.00	1.00	20	3.45	0.94	5	4.20	0.45	5	3.80	1.10
T1	12	3.33	0.49	4	4.00	0.82	13	3.31	0.95	5	3.80	1.10	5	4.00	0.71
T2	11	3.27	0.79	2	3.50	0.71	19	2.32	1.06	5	4.00	1.00	4	3.50	1.00
T3	11	3.00	0.89	2	4.50	0.71	19	2.11	0.99	5	4.20	0.84	4	3.75	0.50
T4	12	4.25	0.75	4	4.25	0.50	13	2.85	1.41	5	4.20	0.84	5	4.20	0.84
T5	12	4.25	0.75	4	4.50	0.58	13	3.00	1.29	5	4.40	0.55	5	3.80	0.84
TX	12	2.92	0.51	4	3.50	0.58	22	2.14	1.04	5	3.40	0.55	5	4.00	0.00
TG1	12	3.83	0.83	4	4.25	0.50	13	2.77	0.93	5	3.80	1.10	5	3.80	0.84
TG2	12	3.67	0.98	4	4.25	0.50	13	2.85	0.90	5	4.40	0.89	5	4.40	0.89
ML1	12	4.25	0.75	4	4.50	0.58	14	3.50	0.85	5	4.60	0.55	3	3.33	0.58
ML2	12	3.75	1.06	4	4.75	0.50	14	3.50	0.65	5	3.60	1.14	3	3.67	0.58
ML3	12	4.17	0.39	4	4.50	0.58	14	3.50	0.76	5	4.20	0.84	3	3.33	0.58
LX	12	2.83	0.39	4	2.75	0.50	22	1.73	1.32	5	3.00	0.00	5	1.60	1.52
ML5	12	3.50	1.17	4	4.25	0.50	14	3.29	0.83	5	4.00	1.22	3	2.67	0.58
MR7	12	3.25	1.14	4	4.25	0.50	22	3.41	1.01	5	4.20	0.45	5	2.80	1.30
ML8	12	4.00	0.95	4	4.25	0.50	14	3.50	0.65	5	4.40	0.55	3	2.67	0.58
MR9	12	3.75	0.97	4	4.50	0.58	22	3.41	1.05	5	4.40	0.55	5	3.20	1.30
MG1	12	3.08	1.08	4	4.25	0.50	22	2.77	0.97	5	4.20	0.45	5	2.80	1.64
MG2	12	3.75	1.06	4	4.25	0.50	22	3.18	1.05	5	4.20	0.45	5	3.60	1.52
CK1	12	1.92	0.90	4	1.75	0.96	22	1.27	0.63	5	2.60	0.55	5	1.20	0.45
CK2	12	2.50	2.28	2	4.50	0.71	16	1.06	1.95	5	4.20	0.45	3	1.33	2.31
CK3	12	2.42	2.23	2	4.50	0.71	16	1.06	1.95	5	4.00	0.71	3	1.33	2.31
CS4	12	3.25	1.14	4	4.00	0.82	22	2.77	1.19	5	4.20	0.84	5	3.00	1.00
CS5	12	4.25	0.87	4	4.25	0.96	22	3.45	1.06	5	4.60	0.55	5	3.80	1.10
CS6	12	3.25	1.06	4	3.25	0.96	22	2.55	0.80	5	4.00	0.71	5	2.60	1.34
CS7	12	3.75	1.22	4	4.25	0.96	22	2.64	0.90	5	4.00	0.71	5	4.00	1.73
CG1	12	3.83	0.83	4	4.25	0.50	22	2.45	1.06	5	4.40	0.55	5	3.80	1.64
CG2	12	4.00	0.74	4	4.25	0.50	22	2.77	1.07	5	4.40	0.55	5	3.80	1.64