The Impact of Instructional Design in a Case-Based, Computer-Assisted Instruction Module on Learning Liver Pathology in a Medical School Pathology Course

By Patricia S. Latham

B.S. in Microbiology, June 1968, Simmons College
M.D. in Medicine, June 1972, University of Southern California

A Dissertation Submitted to

The Graduate School of Education and Human Development
of The George Washington University
in partial fulfillment of the requirements
for the degree of Doctor of Education

January 31, 2015

Dissertation directed by

Maria Cseh
Associate Professor of Human and Organizational Learning

The Graduate School of Education and Human Development of The George Washington University certifies that Patricia S. Latham has passed the Final Examination for the degree of Doctor of Education as of December 12, 2014. This is the final and approved form of the dissertation.

The Impact of Instructional Design in a Case-Based, Computer-Assisted Instruction Module on Learning Liver Pathology in a Medical School Pathology Course

Patricia S. Latham

Dissertation Research Committee

Maria Cseh, Associate Professor of Human and Organizational Learning,
Dissertation Director

Celina Byers, former Associate Professor of Instructional Technology,
Bloomsburg University of Pennsylvania, Consultant Wells Fargo,
Committee Member

Catherine Turley, Associate Professor of Clinical Management and Leadership,
Committee Member

Dedication

This dissertation is dedicated to The George Washington University School of Medicine and Health Sciences Medical School Class of 2015, who were second-year students in the medical school Pathology course in the spring of 2013. I am grateful for the patience and enthusiasm of the students in this class who participated in the studies reported here and who graciously allowed me to report the results of their efforts as the basis of this dissertation. It is a special pleasure to share in the educational process and to learn as much or more than I teach from my students.

Acknowledgements

I would like to especially thank the Advisor for my doctorate, Dr. Maria Cseh, for her guidance and for her unfailing encouragement, support and wonderful enthusiasm for the project and the process. I am also grateful for the efforts of my doctoral committee members, Dr. Catherine Turley and Dr. Celina Byers, who gave their time and experience to guide and help to shape the project from proposal to completion, and to Dr. Brandi Weiss and Dr. Marie Borum who served on my dissertation committee as outside examiners and who provided insightful feedback and strengthened the final dissertation manuscript. I am also grateful to the Medical Faculty Associates, Inc. for tuition funding and to all of my colleagues and students who give of their time and energy to support education in Pathology.

Abstract

The Impact of Instructional Design in a Case-Based, Computer-Assisted Instruction Module on Learning Liver Pathology in a Medical School Pathology Course

The purpose of this quantitative experimental study was to test the impact of three learning interventions on student learning and satisfaction when the interventions were embedded in the instructional design of case-based, Computer-Assisted Instruction (CAI) modules for learning liver pathology in an in-class, self-study, laboratory exercise during a Year-2 medical school Pathology course. The hypothesis was that inclusion of the learning interventions would enhance student satisfaction in using the CAI and improve subsequent CAI-directed exam performance.

Three learning interventions were studied, including the use of microscopic virtual slides instead of only static images, the use of interactive image annotations instead of only still annotations, and the use of guiding questions before presenting new information. Students were randomly assigned to with one of eight CAI learning modules configured to control for each of the three learning interventions. Effectiveness of the CAI for student learning was assessed by student performance on questions included in subsequent CAI-directed exams in a pretest and on posttests immediately after the lab exercise, at two weeks and two months. Student satisfaction and perceived learning was assessed by a student survey.

Results showed that the learning interventions did not improve subsequent student exam performance, although satisfaction and perceived learning with use of the CAI learning modules was enhanced. Student class rank was evaluated to determine if the learning interventions might have a differential effect based on class rank, but there were no significant differences. Class rank at the time of the lab exercise was itself the strongest predictor of exam performance.

The findings suggest that the addition of virtual slides, interactive annotations and guiding questions as learning interventions in self-study, case-based CAI for learning liver pathology in a medical class room setting are not likely to increase performance on subsequent MCQ-based exams, but student satisfaction with use of the CAI can be enhanced, which could provide to be an incentive for students to use similar CAI learning modules for future self-directed learning.

Table of Contents

PAGE

PAGE

PAGE

List of Figures

PAGE

List of Tables

PAGE

.

CHAPTER 1- INTRODUCTION

Computer-assisted instruction (CAI) is increasingly used in medical education as a cost-effective means of decreasing lecture time and providing a multimedia format that can bring relevant subject matter to life in a reusable, cost-effective manner for any number of learners (Berman, Fall, Maloney, & Levine, 2008; Cook, Levinson, & Garside, 2010). Other investigators have found large positive effects on student learning when internet-based instruction is compared to no intervention, but variable and small effects compared to traditional methods, suggesting to several educators that research concerning CAI needs to focus not on comparing it to other methods but on exploring how best to implement the most effective use of it in medical education (Berman et al., 2008; Cook et al., 2008; Cook, 2009). There is a "need to clarify how and when to use e-learning through 'basic science' research and 'field tests' comparing one e-learning intervention to another" (Cook, 2009).

The purpose of this quantitative experimental study was (1) to test the impact on student learning of three instructional design interventions, interactive annotations, guiding questions and the use of whole-digitized virtual slides, on the conceptual and visual learning of liver pathology in the configuration of a Computer-Assisted Instruction (CAI) program used as an enhancement in a Year-2 medical school Pathology course; and, (2) to determine student perceived learning and satisfaction in the use of the CAI program.

Statement of the Problem

CAI are increasingly used in medical education as part of a wave of curricular reform that decreases lecture time and increases the integration of disciplines (Cooke, Irby, & O'Brien, 2010; Cook et al., 2010b; Cook et al., 2010). CAI are often used, however, without careful attention to development or a clear understanding of the instructional design required for optimal learning (Berman et al., 2008; Cook et al., 2008; Cook, 2009; Cook et al., 2010a; Cook, D. A., Levinson, A. J., Garside, S., Dupras, D. M., Erwin, P. J., and Montori, V. M., 2010b; Jha & Duffy, 2002). The Association of American Medical Colleges' Institute for Improving Medical Education reported in 2007 that research was needed to study the effectiveness of interventions in multimedia programs designed for medical education (Association of American Medical Colleges, 2007). In a systematic review of 201 studies comparing the effect of internet-based instruction to no intervention or non-internet interventions for learners in the health professions, Cook et al. (2008) found large positive effects of internet-based instruction compared to no intervention, but variable and small effects compared to traditional methods. Other reviews of CAI effectiveness have arrived at similar findings (Chumley-Jones, Dobbie, & Alford, 2002; Greenhalgh, 2001), suggesting to several educators that research concerning CAI needs to focus not on comparing it to other methods but on exploring how best to implement the most effective use of it in medical education (Berman et al., 2008; Cook, 2009). Cook (Cook, 2009) advocates for the “need to clarify how and when to use e-learning through 'basic science' research and 'field tests' comparing one e-learning intervention to another”.

Research on the instructional design of e-learning materials and CAI for learning pathology is especially limited, even as the use of digitized whole slide images, or virtual slides (VS), has become more common in medical schools as the curricula become more integrated (Pantanowitz, 2012). Pathology courses have moved away from microscopes and the use of glass slides to the use of VS that are accessed by students online. When compared to glass slides and microscopes for reviewing pathology, several studies have shown that students accept VS for learning and appreciate the ease of use, efficiency of study, and ability to access slides from any computer (Kim et al., 2008; Krippendorf & Lough, 2005; Kumar et al., 2004; Nivala, Saljo, Rystedt, Kronqvist, & Lehtinen, 2012). However, only one reported study was found in the literature that compared student exam performance after learning with glass slides or VS, and this study showed that performance with VS was identical or minimally improved compared to historical controls (Kumar et al., 2004). No reported studies have compared the learning of pathology with virtual slides vs static images. In light of the pressing need to know how best to design, administer and assess CAI programs for optimal student learning and, in particular, how to incorporate the use of virtual slides in CAI intended for learning pathology, the following study was conducted.

Purpose of the Study

The purpose of this quantitative experimental study was (1) to test the impact on student learning of three instructional design interventions in the configuration of a Computer-Assisted Instruction (CAI) program used as an enhancement in a Year-2 medical school Pathology course; and, (2) to determine student perceived learning and satisfaction in the use of the CAI program.

Research Questions and Hypotheses

In order to fulfill the research purpose, the following research questions (R) and hypotheses (H) were investigated:

1. R1- Do students achieve greater applied knowledge of the pathology of a disease when they learn features of the pathology in the context of a real tissue section (Virtual Slide) vs static images?

   H1- use of virtual slides (VS) to illustrate pathology for student learning will show a positive correlation with student satisfaction, perceived learning and exam performance

2. R2- Do students achieve greater applied knowledge of the pathology of a disease when illustrations of pathology involve interactivity with associated annotations versus static annotations.

   H2- use of interactive annotations in multimedia will show a positive correlation with student satisfaction, perceived learning and exam performance

3. R3- Do students achieve greater applied knowledge of the pathology of a disease when information is preceded by an introductory question to the information to be learned vs un-cued presentations of information

   H3- use of introductory questions to new information will show a positive correlation with student satisfaction, perceived learning and exam performance

4. R4- Which learning interventions are preferred by students and what features do they attribute to their satisfaction and perceived increase of learning?

   H4- student satisfaction and perceived learning will show a positive correlation with features that give students control of the learning pace, interactivity with the program, ease of navigation and strategies to reduce cognitive load – cued annotations, guiding questions and self-selection of informational content through hyperlinks.

Significance

This study is a direct response to the call for research on principles of multimedia design in medical education (Association of American Medical Colleges, 2007); and in particular, the need to focus research questions on the effectiveness of interventions in design rather than comparisons of Computer-Assisted Instruction (CAI) programs with other modalities of education (Cook, 2009; Berman et al., 2008). The outcome of the interventions described in this proposal provide important information about the instructional design of CAI self-study learning modules that is supportive to the learning of pathology in undergraduate medical students as part of a Pathology course. The information can guide the development of future CAI for learning pathology, when it is used as an enhancement to other educational materials in undergraduate medical education. The need for the information is pressing, since many medical schools, including The George Washington University School of Medicine and Health Sciences (GWU-SMHS), are undergoing medical curriculum reform, with plans for decreased Pathology class time and increased use of self-study programs such as the CAI that is the subject of research in this proposal.

Theoretical Foundation

The use of Computer-Assisted Instruction (CAI) for student learning in medical education draws from theories of adult learning and the development of relevant, case-based, professional expertise on the one hand, and theories of multimedia instructional design and metacognitive strategies to support learning in that medium on the other (see Figure 1 - Conceptual Framework).

The use of a CAI module for learning pathology is based on the multimedia learning theory of Mayer (Mayer, 2001), supported by the cognitive load theory of Sweller (Sweller, 1988). The use of CAI also draw heavily from theories of adult learning based on an appreciation that adults are most comfortable with learning in which they are active participants with a measure of control over their learning (Knowles, Holton, & Swanson, 1998). The use of medical case studies as a basis for the CAI draws on theories of relevance to professional goals as a motivation for learning (Aamodt & Plaza, 1994; Barrows, 1994; Bruner, 1960). Application of basic knowledge to solving case-related problems develops a higher order of learning (Bloom, 1994). The use of relevant case studies also builds on theories of situated learning and on iterative learning and problem solving (Argyris & Schon, 1974; Argyris & Schon, 1996; Kolb, 1984; Lave & Wenger, 1991). CAI can also be used to increase student exposure to cases that would be difficult to see otherwise, allowing them to apply knowledge learned in one case to solve another (Gentner, Loewenstein, & Thompson, 2003).

The instructional design of CAI draws on the theories espoused for learning in this medium. The ADDIE model for CAI development described steps to analyze the specific needs, design the CAI, develop the program, implement and evaluate the results (Malachowski, 2002). Gagne et al (Gagne, Briggs, & Wager, 1992) described specific steps in the development of CAI to promote effective learning. According to the generative theory of multimedia learning, the learner must select from information presented in different modalities (text, image, animation, etc.) that which is important and then organize and integrate it with pre-existing information into new knowledge (Mayer, 1997). The use of interventions in CAI supports interactivity (VS, interactive annotations) that can reinforce learning and improve performance on exams, building on theories of iterative learning and problem solving (Argyris & Schon, 1974; Kolb, 1984). The use of guiding questions supports student reflection and self-explanations that will be help in reducing cognitive load and integrating new information with existing information to create a new understanding that can deepen learning and strengthen long-term memory (Bude, van de Wiel, Imbos, & Berger, 2012; Kirschner, Sweller, & Clark, 2006). The use of hyperlinks generates a higher level of interactivity that can be associated with deeper learning (Ross & Tuovinen, 2001). The use of hyperlinks fosters cognitive flexibility and the appreciation that higher level learning is not linear, but occurs in a matrix of information (Graddy, 2001; Spiro, Feltovich, Jacobson, & Coulson, 1992).

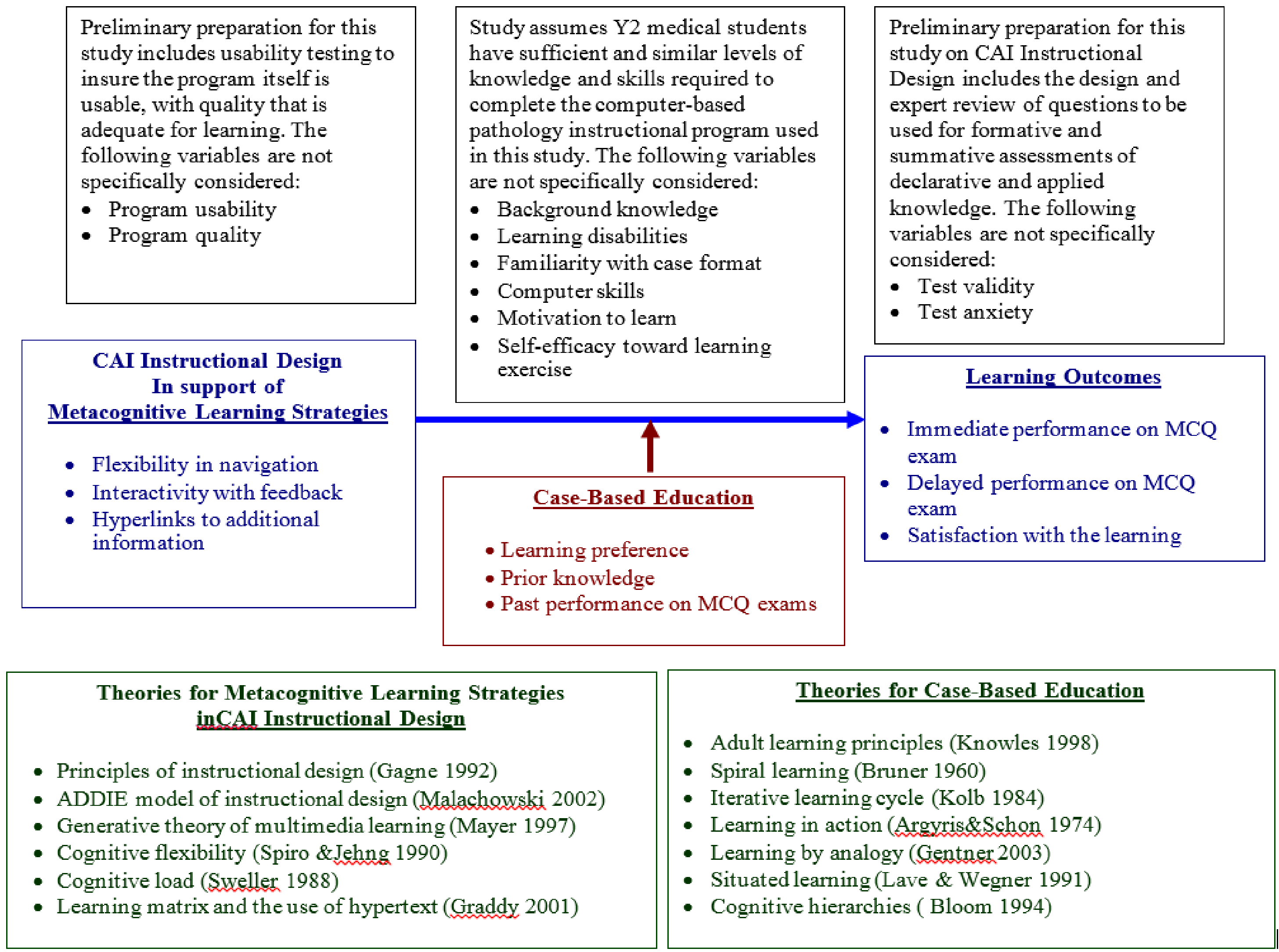


Figure 1. Conceptual Framework

Methodology

This study is an experimental quantitative research study that uses randomization of subjects to experimental and control groups to address research questions and to test hypotheses. The study is considered from the perspective of a positivist epistemology and a realist ontology (Creswell, 2003; Crotty, 2005).

The study is designed to determine the effects of instructional design in a case-based, Computer-Assisted Instruction (CAI) program on student learning as measured by exam performance and student perceptions of learning and satisfaction. The CAI is presented as a laboratory exercise, an enhancement to learning liver pathology in a Year-2 medical school Pathology course at The George Washington School of Medicine and Health Sciences (GWU-SMHS).

The CAI was administered after basic lectures on liver disease in the Pathology course, a one-hour lecture given in class on the Friday before the Monday CAI lab exercise titled Introduction to the Laboratory Diagnosis of Liver Disease, and two one-hour lectures on acute and chronic liver disease that were presented as self-study pre-recoded lectures online. A pre-test was administered prior to the CAI to assure that all students had the same base of knowledge about liver pathology. The pre-test associated with the CAI accounted for 2% of a student's score on the following practical exam that covered the material in the section.

Three variables were studied as interventions in the instructional design of the CAI, resulting in eight CAI modules in which one variable was studied while the other variables were controlled. All students at the GWU-SMHS who registered for the Pathology course in the spring of 2013 and who consented to use of their results for research analysis were participants (n=173 from a total of 176 students), with random assignment of students to one of the eight CAI modules (n=11 per CAI module X2). Each CAI module included a control case study (Case A) and two cases that were configured to test one or more of the interventions, all of the interventions or none of the interventions (Cases B and C),. Lectures were provided in advance to insure that all students had access to the basic information concerning liver disease that was required for application of that information to the case studies presented in the CAI modules, assured by pre-test. Learning that resulted from use of the eight CAI modules was measured by exam performance on a post-test at completion of the laboratory exercise (PostTest 1), at 2 weeks (PostTest 2) and at 2 months (PostTest 3). The PreTest and each PostTest accounted for 0.5% of a student's grade in the Pathology course (2% total). Questions on the two case studies in CAI configured with interventions were mixed in with questions on the control case study, but the responses did not count toward a student's grade in the course. The satisfaction and perceived learning of students using the CAI was assessed by an anonymous survey linked only to the Group number (1-8), with responses recorded on a Likert-type scale and as responses to open-ended questions.

The variables included (1) use of whole slide images as virtual slides (VS) vs static images, (2) use of interactive annotations vs static annotations, and (3) the use of guiding questions. The content of the CAI was the same for each CAI module, including textual descriptions of the images and information about the pathology of the selected diseases presented and hyperlinks to normal or basic information and hyperlinks to additional or advanced information about the selected diseases presented in the CAI.

Assumption and Delimitations

This research to determine the effectiveness of the described interventions in Computer-Assisted Instruction (CAI) programs was limited to the study of Year-2 medical students in mid-session of a year-long medical school Pathology course in which the CAI was used as an enhancement to self-study and in-class lecture materials that presented basic information about the liver diseases and pathology that were included in the CAI. A pretest was administered in class to provide assurance that students had learned the basic information about the diseases presented in the CAI and were prepared to apply that knowledge to learning from the case studies. The study assumed that the medical students had a similar background and experience in using CAI and the use of VS, which were used regularly in teaching pathology within the course. The CAI was a required lab exercise and was administered in-class, with each student accessing their assigned CAI learning module and working independently during the exercise. The CAI was immediately followed by a posttest and satisfaction survey.

Any adverse effects of the controlled conditions were minimized as much as possible by allowing adequate time for self-study of the lecture materials and manual before the pretest; confining the time of the activity to scheduled class time such that it did not impact on other learning activities; designing the CAI and tests in a way that allowed adequate time for completion by most students, as determined by prior usability testing; providing a computer if necessary, adequate work space ; and, the use of a posttest to fairly assess applied knowledge after use of the CAI, as determined by expert review and critique of the questions prior to the exercise.

CHAPTER 2: LITERATURE REVIEW

The literature review was undertaken to determine how Computer-Assisted Instruction (CAI) programs have been used in case-based, multimedia educational programs for medical student learning (see Appendix A, Literature Map) to determine what theories have supported their development and instructional designs, and which designs have proven to be most effective for student learning, with particular attention to strategies for improving the learning of pathology.

Purposes and Methods of Literature Review

References were identified using databases on computers, journals, books and internet sites. Textbooks that were also identified (Mayer, 2001; Plass, Moreno, & Brunken, 2010; Gardenfors & Johansson, 2005). These textbooks are compilations of the work and insights from guiding experts in the field of instructional design in the use of computer-based or web-based programs for education. Databases were searched for articles in English from 1990- present, in order to capture those that focus on programs that are computer-based in nature. Additional articles reporting on theories in use and observations derived from research were selected from the citations within these articles. Databases surveyed include Ovid Medline, ABI/Inform, Academic Search Premier, ERIC, Psychlnfo, Proquest and CINAHL. Additional references were obtained from citations within reviewed publications and also through informal references to selected works accessed online, using search engines such as Google. Search terms alone and in combination included: medical education, multimedia, adult learning, learning, education, multimedia instruction, computer-assisted instruction, computer-aided instruction, computer-assisted design, instructional design and multimedia design, hyperlinks, cueing, Socratic method and cognitive load.

The following themes emerged from the review of the literature: the importance of CAI development and instructional design for student learning; the utility of CAI for case-based learning in medical education; the use of multimedia in CAI and considerations of cognitive load; and guidance within CAI instructional design.

CAI development and instructional design for student learning

Computer-based, multimedia educational programs are often used in medical education as a means of decreasing in-class lecture time and increasing integration of disciplines, but without careful attention to development or a clear understanding of the instructional design required for optimal learning (Berman et al., 2008; Chumley-Jones et al., 2002; Cook et al., 2008; Cook, 2009; Cook et al., 2010; Cook et al., 2010; Greenhalgh, 2001). The Association of American Medical Colleges' Institute for Improving Medical Education reported in 2007 that research was needed to study the effectiveness of interventions in multimedia programs designed for medical education (Association of American Medical Colleges, 2007).

Theories to support the use of computer-assisted instruction as a method of learning in medical education are based on adult learning principles (Knowles et al., 1998). Students have the potential to study at their convenience and to pause the learning or review passages, which allows them to segment the learning as a means to help in controlling cognitive load (Sweller, Merrienboer, & Paas, 1998). In the development of CAI, The ADDIE model (Analysis, Design, Development, Implementation and Evaluation) is a generic model that delineates the sequential components in the process of instructional design (Malachowski, 2002) . Gagne et al. (1992) also proposed principles of instructional design that include nine sequential events of instruction for optimal learning. The instructional system should: (1) gain attention (2) inform learners of objectives, (3) stimulate recall prior learning, (4) present stimulus, (5) provide learning guidance, (6) elicit performance, (7) provide feedback, (8) assess performance, (9) enhance retention and transfer (Gagne et al., 1992). Thus, CAI has the potential to take advantage of many features that theory indicates are important to learning (Berman et al., 2008).

CAI can take advantage of multimedia, such as images, animations and audio as a means of bringing relevant subject matter to life while delivering content in a reusable, cost-effective manner for any number of learners (Berman et al., 2008). Students can learn independently at their own pace and convenience, which enhances satisfaction with this method of learning (Blake, 2010). Berman et al. (2008) point out that CAI are particularly well suited to practice-based learning in medicine, since teaching modules can present patient simulations that situate the lessons in a real world context for students. Assessments can be added to CAI teaching modules to document participation and knowledge gained (Berman et al., 2008). The potential benefits of well-designed computer-based, case-based educational programs in medicine depend, however, on their ease of use and the quality of their instructional design (Blake, 2010). In a systematic review of 201 studies comparing the effect of internet-based instruction to no intervention or non-internet interventions for learners in the health professions, Cook et al. (2008) found large positive effects of internet-based instruction compared to no intervention, but variable and small effects compared to traditional methods. Other reviews of CAI effectiveness have arrived at similar findings (Greenhalgh, 2001; Chumley-Jones et al., 2002), suggesting to several educators that research concerning CAI needs to focus not on comparing it to other methods but on exploring how best to implement the most effective use of it in medical education (Berman et al., 2008; Cook, 2009). Cook (2009) advocates for the “need to clarify how and when to use e-learning through 'basic science' research and 'field tests' comparing one e-learning intervention to another”.

Research on the instructional design of e-learning materials and CAI for learning pathology is especially limited, even as VS have become common in medical schools as the curricula become more integrated (Pantanowitz, 2012). Pathology courses have moved away from microscopes and the use of glass slides to the use of whole-digitized virtual slides (VS) that are accessed by students on line. When compared to glass slides and microscopes for reviewing pathology, several studies have shown that students accept VS for learning and appreciate the ease of use, efficiency of study, and ability to access slides from any computer (Kim et al., 2008; Krippendorf & Lough, 2005; Kumar et al., 2004; Nivala et al., 2012; Pantanowitz, 2012). Only one reported study was found in the literature that compared student exam performance after learning with glass slides or VS, and this study showed that performance with VS was identical or minimally improved compared to historical controls (Kumar et al., 2004). However, no reported studies have compared the learning of pathology with virtual slides vs static images.

CAI programs for case-based learning in medical education

Theories to support the use of case-based learning and other metacognitive learning strategies in medical education consider the use of a medical case as a means of incorporating practical experience into the learning process in a relevant context, a key feature of adult learning principles as described by Knowles et al. (1998). Case-based learning situates the learning within an authentic context that allows students to see the relevance of their learning and further to identify with the social context of the situation (Lave & Wenger, 1991). Application of knowledge in the context of a case builds on Bloom's taxonomy of cognitive hierarchies in which he describes increasing complexity in learning through six levels of activity that represent the problem solving required in case-based learning (Bloom, 1994). Bruner's cognitive and constructivistic theories in education support the use of case-based learning as a means of putting the problem for learning into a context in which learners can apply previous knowledge and experience to find gaps that can be filled in to create a new level of understanding, so-called spiral learning (Bruner, 1960). Iterative learning is also the method of learning stressed by Kolb's learning cycle in which an experience leads to observations and reflection, guiding to conceptualization and generalization, which leads to the testing of implications (experimentation) (Kolb, 1984). Gentner et al. (2003) proposes the theory that analogy in learning provides a structure map, similarities in problems or cases can facilitate learning. Argyris and Schon (Argyris & Schon, 1974; Argyris & Schon, 1978; Argyris & Schon, 1996; Schon, 1983) reinforce the need for reflection in and on action as an essential step in the process of learning suggested by Kolb (1984). Problem-solving and "double-loop learning" are often the methods by which professionals learn by apprenticeship (Argyris & Schon, 1974; Schon, 1983). In medical education, the incorporation of these theories into a model for professional education frequently involves case-based or problem-based learning (PBL) (Barrows, 1994; Savery & Duffy, 1995). A "case" can be considered to be a "problem situation" in case-based reasoning (Aamodt & Plaza, 1994). Spiro and Jehng (1990) propose the theory of cognitive flexibility. This theory has relevance to case-based learning, because it proposes that the learning process is not linear, but often requires a creative recombination of knowledge (Spiro & Jehng, 1990; Spiro et al., 1992).

Multimedia in CAI programs and cognitive load

Cognitive load becomes an important consideration in the instructional design of CAI (Mayer & Moreno, 2003; Mayer, 2005; Sweller et al., 1998; Van Merrienboer & Sweller, 2010; Issa et al., 2011). CAI can use the features of multimedia to bring the lessons to be learned into a real-world context, promoting learning by enhancing the relevance of the lessons to the student. Students have greater control of their learning using CAI, since students can use the program where and when they choose to do so. CAIs also allow students to control the pace of the program as needed for their learning, to stop and start or to redo the program at will. These features support individualized learning and the ability to 'segment' the material to be learned into smaller chunks that are more manageable for learning, allowing the student time to reflect and to integrate new information into a coherent 'schema' (Sweller, 1988; Sweller et al., 1998). It is thought that the construction of such schemata facilitates the transfer of knowledge to long-term memory and deeper learning and enhances the ability to recall and to apply the new information (Van Merrienboer & Sweller, 2010).

CAI can use static or virtual images, diagrams, animations, audio or video in ways that better explain difficult concepts to students or that help students to integrate the information. The use of more than one modality to explain a concept, such as visual and audio, can use dual channels for processing the information, creating a reinforcement of learning in working memory in ways that increase associations to prior knowledge, which in turn enhances learning (Mayer, 2001). CAI can also be designed in ways that support interactivity of students with the program. Interactivity of students in their learning helps to guide and focus attention, which gives students a sense of control in their learning and keeps the learning more active than passive, which is suggested to result in greater retention of lessons learned (Mayer, 2001; 2005; Issa et al., 2011).

Theories to support the use of instructional design to adjust cognitive load and to enhance learner control are largely based on the work of Sweller, et al. (Sweller, 1988; Sweller et al., 1998; Van Merrienboer & Sweller, 2010). Sweller (1988) introduced the theory of cognitive load for instructional design in teaching materials and CAI. This theory suggests that prior knowledge determines what the working memory can absorb at one time. Working memory is limited to the information that can be held at one time and is quick to fade, but long-term memory has permanent or a least indefinite capacity and retention. Cognitive overload refers to the condition in which information presented to the working memory at one time exceeds the capacity for processing it into a complete construct or 'schema' which can be transferred to long-term memory (Sweller et al., 1998). Three types of cognitive load exist (Plass et al., 2010). Intrinsic cognitive load is the interaction of the information (difficulty of topic and volume of information) to be learned with the capability, previous experience and knowledge of the learner to process it; as a rule, cognitive load will be inverse to previous knowledge and experience. The difficulty of a topic often relates to how many disparate bits of information need to be held in the working memory at one time and then processed into a complete understanding of the information or schemata (Sweller et al., 1998). Schemata can be considered as "a set of interconnected propositions centering around a general concept, and linked peripherally with other concepts" (Gagne et al., 1992). A schema is then stored in long-term memory and can be used to accommodate new information by adjustments in the existing schemata or it can facilitate learning of new information with a similar scheme, referred to as ‘automation’ (Sweller et al., 1998). The information directly relevant to the mental construction of such schemata is referred to as ‘germane congnitive load’ and information irrelevant to the construction of the schemata is referred to as ‘extraneous’ cognitive load (Sweller, 1988). The Cognitive Load Theory suggests that instructional design can add unneeded complexity to learning (Sweller, 1988). Specific steps suggested to reduce cognitive load in multimedia learning include such methods as the avoidance of extraneous information; the alignment of words and text with visual images in contiguity; and the use of signals or cues that suggest how the information should be processed (Mayer & Moreno, 2003).

Theories to support the use of interactive annotations as a means of enhancing learning and retention in CAI are based on several principles. Points of interactivity between the student and the program enhance student control, a concept of adult learning, and focuses attention on the material to be learned (Gagne et al., 1992; Knowles et al., 1998). In addition, the points of interactivity can enhance cueing to direct the student to the essential information to be learned (Sweller et al., 1998). In consideration of learning pathology, visual cues such as annotations, arrows or labels, can be used in multimedia to decrease cognitive load (Mayer, 2001; Sweller, 1988). Annotations can direct attention by text or audio to the important points or features in a visual image that are required for learning, thus decreasing extraneous cognitive load. When annotations are not included, a ‘split-attention’ effect can exist when the learner needs to pay attention to a text description of an image, for example, and differentiating the described features of the pathology from less important features of the image (Van Merrienboer & Sweller, 2010). Annotations are also more efficient and less time-consuming, providing the student with more time to process the information into working memory (Tabbers, Martens, & Van Merrienboer, 2004). A detractor may be, however, that less time is spent in seeing the pathology illustrated by the annotation in the context of other features apparent in the image, which could inhibit the ability of students to recognize the features in future images of the pathology.

Theories to support the use of hyperlinks are based on theories of cognitive flexibility which propose that information on the computer and web can be linked together (hypertext) in an interlocking matrix (Graddy, 2001). Hyperlinks can support deeper learning in web-based programs (Ross & Tuovinen, 2001). Graddy (2001) proposed that each level of knowledge exists in an inter-relationship with the other information to form an overall matrix of information. The information provided in the hyperlinks must be tailored, however, to the knowledge-level of the learner.

Role of Guidance within CAI program instructional design

An emerging concern in the instructional design of CAI is whether or not the student should be guided in his or her learning or left to construct their own new knowledge when provided with the resources (Bude et al., 2012; Kirschner et al., 2006). Theories to support the use of guiding questions as a means of enhancing learning and retention in CAI include the ability of a guiding question to draw attention to the material to be learned, but also to promote 'self-explanations' that stimulate reflection and reasoning for one's self, a process that increases cognitive activity (Bude et al., 2012; Chi, Bassock, Lewis, Reimann, & Glaser, 2013). At the same time, the strategy is one that tends to decrease cognitive load by helping the student to consider the more specific information that is needed for learning.

> Thalheimer (2003) suggests that "Well-designed questions are particularly effective because they (1) provide learners with practice retrieving information from memory, (2) give learners feedback about their misconceptions, (3) focus learners' attention on the most important learning material, and (4) repeat core concepts, giving learners a second chance to learn, relearn, or reinforce what they previously learned or tried to learn" (p. 3)… questions act as cues to trigger memory searches, and ultimately, recall of the appropriate thoughts and answers" (p. 8) .

Guiding questions help novice students, in particular, to accumulate and process information more quickly in working memory, increasing the likelihood that it will be retained in long-term memory (Bude et al., 2012; Kirschner et al., 2006; Mayer, 2004).

> Krischner (2006) makes the point that "empirical studies over the past half-century that consistently indicate that minimally guided instruction is less effective and less efficient than instructional approaches that place a strong emphasis on guidance of the student learning process. The advantage of guidance begins to recede only when learners have sufficiently high prior knowledge to provide "internal" guidance". (p. 1)

Inferences for Research Study

The literature indicates that there is a need for determining optimal instructional design for Computer-Assisted Instruction (CAI) programs developed for medical student education (Cook et al., 2010). The use of case-based CAI has support in learning about medical diseases both from the standpoint of adult learning principles and from the development of professional expertise by working through examples of problems that will be reiterated frequently in practice (Argyris & Schon, 1974; Lave & Wenger, 1991; Schon, 1983). The creation of interactivity in the instructional design of CAI is an important means of commanding attention and providing a sense of control to students (Mayer, 2001). The use of virtual slides in pathology can help to meet that need and also puts the case in a context that is more real for the learner (Kim et al., 2008; Kumar et al., 2004; Nivala et al., 2012). The use of annotations draws the attention of the student to important points for learning, but interactivity with the annotation can provide an opportunity for reflection and self-explanation that may deepen learning (Tabbers et al., 2004; Van Merrienboer & Sweller, 2010). The use of guiding questions may also draw attention to learning points and guide instruction, as well as providing another opportunity for self- explanation to deepen learning, especially for the novice learner (Bude et al., 2012; Kirschner et al., 2006). These interventions that draw attention to the points of learning help to decrease extraneous cognitive load (Mayer, 2005; Sweller et al., 1998; Van Merrienboer & Sweller, 2010).

CHAPTER 3: METHODOLOGY

The purpose of this experimental study was (1) to test the impact of three instructional design interventions in the configuration of a Computer-Assisted Instruction (CAI) program on student learning when used as an enhancement in a Year-2 medical school Pathology course; and, (2) to determine student perceived learning and satisfaction in the use of the CAI program.

In order to fulfill the research purpose, the following research questions (R) and hypotheses (H) were investigated:

1. R1- Do students achieve greater applied knowledge of the pathology of a disease when they learn features of the pathology in the context of a real tissue section (Virtual Slide) vs static images?

H1- use of virtual slides (VS) to illustrate pathology for student learning will be positively correlated with student perceived learning and exam performance

2. R2- Do students achieve greater applied knowledge of the pathology of a disease when illustrations of pathology involve interactivity with associated annotations versus static annotations.

H2- use of interactive annotations in multimedia will be positively correlated with student perceived learning and exam performance

3. R3- Do students achieve greater applied knowledge of the pathology of a disease when information is preceded by an introductory question to the information to be learned vs un-cued presentations of information

H3- use of introductory questions to new information will be positively correlated with student perceived learning and exam performance

4. R4- Which learning interventions are preferred by students and what features do they attribute to their satisfaction and perceived increase of learning?

H4- student satisfaction and perceived learning will be positively correlated with features that give students control of the learning pace, interactivity with the program, ease of navigation and strategies to reduce cognitive load – cued annotations, guiding questions and self-selection of informational content through hyperlinks.

A quantitative, randomized and controlled experimental design was selected as the methodology for testing the hypotheses proposed in this study.

Sample and Population

The sample included all medical students enrolled in the second-year GWU-SMHS Pathology course in 2012-13, who participated in the Computer-Assisted Instruction (CAI) laboratory exercise and consented to research analysis of their results.

A power analysis, using GPower3 (ANOVA main effects and interactions, α error .05, df 1) and assuming an effect size of 0.25 and a power of 0.90, indicated that a total sample size of 171 would be sufficient to determine statistical significance. Thus, the sample size of 173 was considered adequate for the statistical analysis that would be necessary.

The CAI learning module was the method chosen for presenting illustrative cases and pathology of common liver diseases in the 2012-13 academic year. The sample included 173 students from a total of 176 participants (see Table 1).

Table 1

Mean ± SD Age and Gender of Sample and Student Groups

Student Groups	Number of students	Mean Age in Years		Males		Females	
		Mean	SD	N	%	N	%
Group 1	23	24.7	2.1	11	48	12	52
Group 2	22	24.8	2.2	11	50	11	50
Group 3	22	24.7	1.7	8	36	14	64
Group 4	21	25.4	2.8	15	71	6	29
Group 5	21	26.9	3.6	6	29	15	71
Group 6	20	24.9	1.9	16	80	4	20
Group 7	22	25.2	1.8	9	41	13	59
Group 8	22	25.9	2.3	13	59	9	41
Total	173	25.2	2.4	89	51	84	49

The population represented by the sample is second-year medical students in a curriculum similar to that of other medical curricula in the United states, who are learning pathology in the context of an in-class course and who have acquired experience with text-based descriptions of cell and tissue pathology and the interpretation of that pathology demonstrated in static images and by the use of virtual slides. A comparison of student demographics in the sample group from the George Washington University School of Medicine and Health Sciences (GWU-SMHS) and the population of U.S. medical students is presented in Table 2, derived from a publication of the American Association of Medical Colleges in 2012 (Castillo-Paige, 2012) .

Table 2
Percentage of Medical Student Sample at The Geroge Washington University School of Medicine and Health Sciences and Population of Medical Students in the U.S. by Age,

Medical Students	Males %	White %	Black %	Latino or Hispanic	Asian %	Native American** %	Foreign %	Other %
Sample	40	56	12	2	28	1	1	0
Population	52	58	6	8	20	<1	<1	7

Note . Data reported by American Association of Medical Colleges 2012 (Castillo-Paige, 2012).
* Sample on entry to medical school 2011, altered at time of the CAI lab exercise by student transfers in and out of class
** Native American includes Alaskan and Hawaiian students

Experimental Research Design

The methodology selected for this research was an experimental, randomized and controlled study in which all medical students enrolled in the second-year Pathology course and who consented to research analysis of their results (n=173 from a total of 176 students) participated in a laboratory exercise that tested the impact of instructional design of a case-based Computer-Assisted Instruction (CAI) learning module on the effective learning of liver pathology.

Three case studies of liver disease were presented in the CAI learning modules. The variables of instructional design that were tested in the CAI included: (1) interactive VS to illustrate pathology versus static images; (2) interactive annotations of images versus static annotations; and (3) insertion of guiding questions and interactive visualization of answers versus answers presented as static information without guiding questions. Eight CAI learning modules were configured to test and control for each of the three instructional learning interventions (see Table 3). Three case studies were included in each CAI (Cases A-C). Case A was an internal control that was presented in the same way for all students, including static images, static annotations and no guiding questions.

Cases B and C were configured to include the assortment of interventions indicated in the experimental design shown in Table 3.

Table 3

Experimental Design of Learning Interventions in CAI Learning Modules, Case A (without learning interventions) and Cases B-C (with learning interventions)

		CAI Cases		Learning Interventions		
Student Groups	Numbers of students	Case A	Case B-C	Virtual slides	Interactive annotations	Guiding questions
1	23	A	B-C			
2	22	A	B-C			y
3	22	A	B-C		y	
4	21	A	B-C		y	y
5	21	A	B-C	y		
6	20	A	B-C	y		y
7	22	A	B-C	y	y	
8	22	A	B-C	y	y	y

The impact of the learning interventions was measured by performance of student groups on posttest exams as the dependent variables, measured as the percent correct answers on multiple choice questions (MCQ) directed to Case A (without learning iterventions) or to Cases B-C (with learning interventions) on PostTest exams immediately after the CAI lab exercise (PostTest 1), two weeks after the lab exercise (PostTest 2) or two months after the lab exercise (PostTest 3). were the dependent variables used in statistical analyses. Quantitative results of exam performance in the posttest exams counted for 1.5% of a student's overall grade in the Pathology course.

Research Procedures

The exercise occurred in the mid-year of the Pathology course in February 2013, by which time students had acquired a working knowledge of the concepts of general pathology and a familiarity with learning how to apply new information to case-based studies in pathology. All students were provided with background information about liver disease and pathology before working with the CAI module, including a one-hour lecture given in class on the Friday before the Monday CAI lab exercise, entitled Introduction to the Laboratory Diagnosis of Liver Disease, and two one-hour lectures on acute and chronic liver disease that were presented as self-study, pre-recoded lectures online. These materials were provided to the students for self-study at least one week prior to the CAI lab exercise, which was scheduled on a Monday. Pre-recorded lectures included Power Point slides with accompanying audio.

Students had experience with self-study of online lectures before class and case-based exercises during class, since this was the format that was used for learning renal pathology in a previous section. In order to assure that all students began work with the CAI from the same base of information about liver pathology, all students took a closed-book pretest consisting of ten multiple choice questions (MCQ) that tested declarative knowledge about the information presented in the lectures (PreTest). Student scores on the PreTest counted for 0.5% of their grade in the Pathology course.

The lab exercise with the CAI took place in class following the pretest. Students were randomly assigned by Blackboard Class Management software to one of eight student groups. Each student group had access to only one of the eight CAI learning modules configured according to the instructional designs necessary to test the research hypotheses. In order to assess the learning impact of a given instructional design on each student, students worked independently on an assigned CAI module (Groups 1-8) accessed through the Blackboard course management website, insuring that students could only access the CAI modules and hyperlinks individually assigned to them. Students used a lab computer with internet connectivity or their own laptop computer with WiFi connectivity. All of the interactivity designed in the modules was functional for all students, although it was necessary for students using Google Chrome as a browser or using an IOS operating systems to access hyperlinks in a separate file in Blackboard. All students were provided with the same information about each case.

Each CAI module had the same informational content, including: images showing key features of liver pathology on static photographs of pathology shown on the VS or on the VS itself; annotations of key pathologic features of liver disease on static images or on VS; and hyperlinks to basic and advanced additional knowledge about liver pathology. The CAI modules differed, however, in how the content was presented. The experiments were designed to test the hypotheses of the study in the following ways.

Hypothesis (H1) was tested by comparing the results of students using CAI modules with VS of pathology (Groups 5-8) to the results of students using static images (Groups 1-4) in the CAI instructional design. Students using VS had more control and interactivity with the pathology images than students assigned to static images. Students were able to manipulate the VS in a way that showed several examples of the key features of the pathology in the context of other tissue findings and artifacts. Annotations were added to point out the key features on the VS. Students assigned to CAI modules using annotated static images were not directed to use VS, and so, did not have the advantage of seeing multiple examples of key pathologic features or the experience of learning to differentiate key features of pathology from artifacts; instead, their attention was focused only on the essential features of pathology that needed to be learned. These students still had access through hyperlinks to VS that were not annotated, but they did need to select the hyperlink.

Hypothesis (H2) was tested by comparing the results of students using CAI learning modules with image annotations that appeared only when selected (Groups 3,4,7,and 8) to those in which annotations appeared automatically (Groups 1,2, 5, and 6). Students using interactive annotations had more control and greater interactivity with the pathology images and textual information. They had the opportunity to test themselves on recognition of important pathology features in an image and to reflect on important information about a case before selecting the annotated image or textual information. Students assigned to CAI modules using static annotations did not have the interactivity or opportunity to test themselves; instead, their attention was promptly directed to important features of the pathology to be learned about the case.

Hypothesis (H3) was tested by comparing the results of students using CAI learning modules with interactive, introductory guiding questions to important information presented about the cases (Groups 2,4, 6, and 8) or no guiding questions (Groups 1,3, 5, and 7). Students using interactive guiding questions had more control and greater interactivity with important information presented about the cases and they had the opportunity to consider the "answer" before reviewing important information about the case. Students assigned to CAI modules presenting important information without guiding questions did not have the interactivity or opportunity to test themselves; the important information about the case was presented directly and without challenge.

During the time students were working with the CAI, they were allowed to refer to notes about the liver diseases presented in the CAI, similar to the way in which they might work with the CAI at home; however, students were not allowed to collaborate with one another. Students had two hours and 15 minutes to complete the CAI module.

On completion of the CAI learning module, students took a posttest of eleven MCQ that tested applied knowledge concerning the liver pathology presented the CAI modules, three MCQ directed to Case A and eight MCQ directed to Cases B and C (PostTest 1 at Time Point 1). Some of the questions about each case included images. Student scores on MCQ directed to Case A, the internal control case study that all students had in common, counted 0.5% toward their grade in the Pathology course. In order to remove any concern that the learning interventions might advantage or disadvantage a student's grade in the course, MCQ directed toward Cases B and C did not count toward their grade in the course. However, questions directed toward Cases A, B and C were mixed together and students were not told which questions counted toward their grade in the course. Different MCQ of a similar style used to test applied knowledge were also used in later exams. A second exam was given one two weeks later during a sectional Pathology course exam that included other questions on Gastroenterology and the Musculoskeletal systems (PostTest 2 at Time Point 2). A third exam was given two months later and included questions on the Endocrine and Reproductive systems (PostTest 3 at Time Point 3). Similar to the immediate posttest exam, application questions that applied to Case A counted 0.5% toward the student's grade in the course for each exam, but these questions were mixed with questions about B and C, which did not count toward a student's grade.

Hypothesis (H4) that student satisfaction and perceived learning would show a positive correlation with features that gave students control of the learning pace, interactivity with the program, ease of navigation and strategies to reduce cognitive load was tested by the use of an anonymous online survey (Survey Monkey) that included ten questions measuring student perceptions of learning and satisfaction in using the CAI modules on a six-point Likert-type scale and three open-ended questions (see Appendix B). Responses to open-ended questions concerning student reactions to specific attributes in the program was solicited in written form online and analyzed for themes and categories that address strengths and weaknesses in the CAI instructional designs for learning liver pathology.

Usability Testing and Exam Development

CAI modules were developed and tested with the support of Information Technology at GWU to insure that the interactive features and hyperlinks were functional when the modules were accessed through Blackboard, the class management program, using Internet Explorer on classroom PCs with wired internet access.

The case studies with liver pathology and virtual slides that were used in the CAI learning modules were selected from liver cases and virtual slides that had been used in previous years in the Pathology course in a small group setting. Multiple Choice Questions (MCQ) for the Pre-Test were taken in part from exams given to Year-2 medical students in previous years, testing declarative knowledge concerning the basic information presented in the introductory lecture and pre-recorded lecture materials. Other MCQ for the PreTest were developed for the purpose of the lab exercise, also designed to test declarative knowledge concerning the preparatory materials. All MCQs for Post-Tests 1-3 were configured to test applied knowledge. MCQ were written by GWU-SMHS faculty with experience on the Pathology Item Writing committee of the National Board of Medical Examiners (NBME) and using the guidelines prepared by the NBME (Case & Swanson, 1998). Each question was reviewed by faculty at GWU-SMHS with experience in writing questions that test applied knowledge and in a format that avoided pitfalls. Questions were reviewed by two content experts, 3rd or 4th year medical students, Pathology residents or faculty, for accuracy and clarity. These assessments insured that the CAI cases and presentations were easily navigated and user-friendly; the questions were clear and understandable; and, that the medical information was correct.

Data collection

Performance on MCQ exam questions and survey questions were collected on individual scantron sheets and analyzed using the technology programs in place at the George Washington University School of Medicine and Health Sciences (GWU-SMHS), a program that reports performance of each student and of the class as a whole on each question and also the performance of each student and the class as a whole on the entire exam. Data on MCQ exams was collected at three time points after participation in the Computer-Assisted Instruction (CAI) modules, immediately after (PostTest 1 at Time Point one), 2 weeks later during the practical exam for the section on the Pathology course relevant to the CAI modules (PostTest 2 at Time Point two) and 2 months later during the final exam in the course (PostTest 3 at Time Point three).

Student responses to survey questions concerning the CAI learning modules were obtained immediately following the laboratory exercise and PostTest 1 using an anonymous online survey program, Survey Monkey (www.surveymonkey.com). Ten questions measured student perceptions of learning and satisfaction in using the CAI modules on a six-point Likert-type scale ranging from 1 (Strongly Agree) to 6 (Stongly Disagree). Three open-ended questions asked each student their opinions about the CAI learning modules. Results of descriptive statistics were provided by the Survey Monkey program for each Likert-style question by student group. Student comments on open-ended questions were also organized by student group in the results available within the program. Students comments were then organized into categories and themes for reporting of results.

Instruments

The only instrument used in this study was the survey questionnaire evaluating student satisfaction and perceived learning. The questionnaire used was one was developed for evaluation of ‘Reusable Learning Objects” (RLO) Centre for Excellence in Teaching and Learning in Reusable Learning Objects (CETL), available as an evaluation tool for download from the RLO-CETL toolkit (Centre for Excellence in Teaching and Learning in Resuable Learning Objects (RLO-CETL), 2005) and used as presented in a survey evaluating the use of RLO by nursing and medical students (Blake, 2010), although a six-point Likert-type scale rather than a four-point Likert-style scale was used for student evaluation. A 6-point scale was used to provide a range of choice and to force students to select a response at the positive or negative end of the scale. The survey questionnaire used here is an assessment tool consisting of 10 items to which the participant responds on a Likert-type scale, from 1 (Strongly Agree) to 6 (Strongly Disagree) in the student evaluation of the CAI learning modules and three open-ended questions: what was good or helpful; what was bad or distracting; and, what suggestions for improvements. (see Appendix B) .

Data analysis

Variables

Three learning interventions were considered as categorical independent variables for evaluation in the instructional design of the Computer-Assisted Instruction (CAI) learning modules, including: (1) the use of virtual slides (VS) of pathology or static images; (2) image annotations that appeared only when selected instead of appearing automatically; and, (3) guiding questions before delivery of information to be learned. The dependent variables were the student exam scores, reported as percent of correct answers to multiple choice questions (MCQ), on a pretest exam before use of the CAI learning module (PreTest 1) and on three posttest exams occurring just after use of the CAI (PostTest 1), at 2 weeks after use of the CAI (PostTest 2) and at two months after use of the CAI (PostTest 3).

After consideration of the results that emerged from the data after analysis of the original three independent variables, additional studies were added to determine if the learning interventions might have a differential effect on exam performance in students with a high or low background knowledge or aptitude for learning pathology. Kalyuga (2010) reported that hyperlinks could have a differential effect on students with high- or low-level knowledge, so it seemed important to consider the possibility that learning interventions in the instructional design of these CAI learning modules might also have differential effects on students with a lesser or greater aptitude for learning pathology. For the purposes of this study, class rank was determined by a student's cumulative average score in the Pathology course at the time of PostTest 2 (Cum Ave). Statistical analyses were included to test for this variable. (see Table 4)

Statistical Analyses

The statistical tests applied to determine the impact of the learning interventions are summarized in Table 4. The individual learning interventions and combinations of the interventions were also studied as they are expressed in the instructional design of the CAI modules assigned to student groups (Groups 1-8). The predictive ability of the individual learning interventions was studied by multiple regression. An investigation of class rank was added in the course of data analysis and the tests that explore the impact of class rank are included in the list of statistical analyses shown in Table 4. All statistical analyses used IBM SPSS Statistics (Version 22).

Table 4. Summary of Statistical Tests for Analysis of Learning Interventions

Table 4

Statistical Tests Applied To Determine the Impact of Student Group Assignment and Learning Interventions (Virtual Slides, Interactive Annotations and Guiding Questions) on PreTest and PostTest Exam Performance for Multiple Choice Questions (MCQ) Directed to Case A (no learning interventions) or Cases B-C (with learning intervetions), and Added Analyses to Determine the Impact of Class Rank

Study	Measurement	Statistical Test
Effect of Student Group Assignment on Exam Performance	Between Student Group Scores on the PreTest Between Student Group Scores on PostTest 1-3 MCQ Directed to Case A Between Student Group Scores on PostTest 1-3 MCQ Directed to Cases B-C	One-way ANOVA
Effect of Student Group Assignment on Exam Performance (confirmation of ANOVA)	Between Student Group Scores on the PreTest Between Student Group Scores on PostTest 1-3 MCQ Directed to Case A Between Student Group Scores on PostTest 1-3 MCQ Directed to Cases B-C	Kruskall-Wallis
Effect of Student Group Assignment on Exam Performance for MCQ Directed to Case A compared to MCQ Directed to Cases B-C	Within Student Group Test Scores on PostTest 1-3 MCQ Directed to Case A compared to Case B-C	Paired Samples Student t-test
Effect of Student Group Assignment on Sequential PostTest Exam Performance	Within Student Group Sequential Scores on PostTest 1-3 MCQ Directed to Case A Within Student Group Sequential Scores on PostTest 1-3 MCQ Directed to Cases B-C	Repeated Measures ANOVA
Effect of Quintile Class Rank (High, Mid, Low) on Exam Performance	Between Student Quintile Rank Scores on PreTest Between Student Quintile Rank Scores on PostTest 1-3 MCQ Directed to Case A Between Student Quintile Rank Scores on PostTest 1-3 MCQ Directed to Cases B-C	One-Way ANOVA
Correlation of Class Rank as a Continuous Variable and Exam Performance	PreTest Post-Test 1, 2 and 3 MCQ Directed to Case A Post-Test 1, 2 and 3 MCQ Directed to Cases B-C	Pearson correlation correficient
Predictive Ability of the Three Learning Interventions and Class Rank on Scores for MCQ Directed to Cases B-C	Predictive ability of VS, IA, GQ and also High- and Low-Quintile Class Rank on Scores for PostTest 1-3 MCQ directed to Cases B-C Predictive ability of VS, IA, GQ Within High Quintile Student Group Directed to Cases B-C for MCQ Directed to Cases B-C Predictive ability of VS, IA, GQ Within High Quintile Student Group Directed to Cases B-C for MCQ Directed to Cases B-C	Multiple Regression
Effect of Student Group Assignment on Responses to Likert-style Questions in the Survey Questionnaire	Between Student Group Median Scores on 6-point Likert-style Questions	Kruskall-Wallis

Note. VS - Virtual Slides,IA = Interactive Annotations, GQ = Guiding Questions
Results for testing conditions and assumptions for each Statistical Test and Measurement are described in Appendix D

Statistical Conditions and Assumptions

Extreme outliers were determined for the statistical analyses using Grubb's test at an alpha level < .01 (Grubbs, 1969). The normal distributions of exam scores within each group were determined by skewness and kurtosis at alpha level < .01 ($z < \pm 2.58$). The alpha level for all other statistical tests was set at .05. The statistical tests applied are described below, with further description and results of testing for conditions and assumptions in Appendix D.

Analyses by one-way ANOVA used Levene's test to determine Homogeneity of Variance (HOV). The alpha level for the ANOVA result was set at .05. If one-way ANOVA results showed statistical significance, Tukey post-hoc tests were done to determine where exactly the differences were manifest. If Levene's HOV was violated, Welch's ANOVA was used for analysis of statistical significance. If Welch's ANOVA results showed statistical significance, Games-Howell post-hoc tests were done to determine where exactly the differences were manifest. The effect size of results were reported as eta squared (η^2).

Analyses by repeated measures ANOVA included Mauchly's test of sphericity at an alpha level .05. If the assumption of sphericity was violated, a Greenhouse-Geisser correction was applied to determine statistical significance at an alpha level .05. Effect sizes were reported as partial eta squared, ηp^2.

Analyses by multiple regression tested assumptions and conditions to determine: independence of observations (errors or residuals), as determined by the Durbin-Watson statistic with a result close to 2.0; a linear relationship between the predictor variables (and composite) and the dependent variable; the homoscedasticity of residuals (equal error variances); and, the absence of multicollinearity. Significant predictor variables were identified at an alpha level .05. Effect size was determined by the calculation of the adjusted R square (*adj.* R^2). The ANOVA test was used to determine if the independent variables had a statistically significant ability to predict the dependent variable at an alpha level of .05.

Paired-Samples t-test determined statistical significance at an alpha level of .05. Effect sizes were reported as Cohen's *d.*

Pearson's correlation coefficients were determined at an alpha level of .05 and reported as a positive or negative correlation. Effect sizes were reported as Pearson's correlation coefficient, *r*.

Kruskal-Wallis non-parametric testing approximated similar distributions of scores for comparative groups by observation of box-plots. Statistical significance was set at an alpha level of .05. If results were statistically significant, pairwise comparisons were made and a Mann-Whitney U test was used for each pairwise comparison. A Cronbach's alpha coefficient was measured to determine internal consistency of the survey results.

Violations of assumptions are described for each statistical analysis when they occurred, detailed in Appendix D. There were very occasional group scores that showed an outlier, such as the one identified in Group 4 scores for PostTest 1 MCQ directed to Cases B-C. This student score was not deleted, since the outcome was unchanged when the score was eliminated. On the basis of this testing, it was determined that occasional outliers would be left in the analyses. There were also occasional deviations in normality, but these data were also left in the analyses unchanged, since the ANOVA test is relatively robust to deviations in normality, especially when numbers in groups are nearly equal (Hinkle, Wiersma, & Jurs, 2003). However, a Kruskal-Wallis test was also run to confirm that there were no statistically significant differences in test result scores between the groups in the PreTest. and PostTest 1-3 exams. Similar to the ANOVA test, the paired-sample t-test is also considered to be robust to violations of normality (Hinkle et al., 2003).

Analysis of Open-ended Survey Questions

Responses to open-ended questions were analyzed by sorting all statements from each group into categories and themes based on similarities, using the method of compare and contrast (Creswell, 2003). The validity of the identified categories and themes was confirmed by a second reviewer with established experience in this qualitative technique.

Ethical Precautions

The results of exam performance in the study were known to this investigator, since Dr. Latham is Director of the Pathology Course and the exercise using Computer-Assisted Instruction (CAI) was an integral part of the course. However, presentation of results uses only de-identified data. The laboratory exercise covered materials that students need to learn in this section of the Pathology course, so attendance was expected. Post-test questions concerning the CAI modules that were configured to test three variable interventions were presented in a mix of questions with other questions directed toward a control CAI case study that was included in each of the test CAI modules. Only the questions concerning the control CAI case study counted toward a student's grade in the course. This same examination strategy was used in all the post-test exams (Time Points 2 and 3) intended to test learning retention on the case studies presented in the CAI modules. Students also completed a survey for satisfaction and perceived learning in their use of the CAI modules. The survey asked students to indicate which CAI Group (1-8) they were assigned to, but their responses were entered anonymously online using Survey Monkey. They were invited to enter their name, but it was not required. Completion of the survey was considered to be implied consent to use the results.

The office of Human Research Institutional Review Board approved the proposal as an exempt study #011322. Students (n = 173) were provided with an informational sheet describing the nature of the research study and provided written consented to use of their test results for research analysis.

CHAPTER 4: RESULTS

The description of the studies and results are presented in this section. Additional details concerning results and SPSS reports, as well as descriptions and results of testing for conditions and assumptions of statistical testing are described in Appendix D.

Description of the Laboratory Exercise

The impact of three interventions to enhance active learning in the instructional design of Computer-Assisted Instruction (CAI) modules was tested by exam performance and satisfaction survey for 173, Year 2 medical students. The three interventions - the use of virtual slides, the use of interactive annotations and the use of guiding questions, were tested in eight case-based learning modules that were configured to control for each of the interventions (see Methods - Study Design). The CAI modules presented three clinical cases, Cases A, B and C. Case A was presented as an internal control, without interventions, in each of eight CAI modules. Cases B and C were presented with variable interventions in the CAI modules, as outlined in the Study Design. Students were randomized to one of eight groups (Groups), each group assigned to one of the eight CAI modules for self-study during an in-class laboratory exercise.

The pretest (PreTest) measured declarative knowledge and was intended to assure that students were prepared to apply the basic knowledge learned in the preparatory materials to the three clinical cases presented in the CAI learning modules. Tests administered after the CAI lab exercise measured applied knowledge gained in working through the CAI learning modules. Each PostTest included three multiple choice questions (MCQ) directed at Case A, four MCQ directed at Case B and four MCQ directed at Case C. Some of the MCQ were associated with images. One test was administered immediately after the lab exercise (PostTest 1), another at two weeks after the lab exercise (PostTest 2), and a third test at two months after the lab exercise (PostTest 3).

Effects of Student Group Assignment on Exam Performance

Each student group (Groups 1-8) represented a specific configuration of the three learning intervention variables in the CAI learning module. Case A in the CAI was without learning interventions and Cases B-C included the learning interventions.

PreTest

The PreTest of declarative knowledge just prior to the CAI lab exercise had a mean score of 75.7 ± 17.5% for all 173 students (Table 5).

Results for each of the eight student groups on the PreTest were analyzed between groups by one-way ANOVA and there were no statistically significant differences between student groups in PreTest scores, $F(7,165) = .467$, $p = .858$, $\eta^2 = .019$.

A Kruskal-Wallis H test was also run to determine if there were differences in test result scores between the groups in the PreTest. Median test scores were not statistically significantly different between groups on the test, PreTest result X2(7) = 3.031, $p = .882$

The results indicate that students acquired sufficient knowledge during their preparations for the lab exercise to achieve at least an average score on the PreTest. The results also indicate that the performance of the student groups on the PreTest were statistically equivalent to one another, suggesting that the randomized assignment of students to the groups was able to create groups of students with similar knowledge of the subject matter at the time of the PreTest.

Table 5

Mean % ± SD Between Student Group Scores on the PreTest and on PostTest 1-3 on Multiple Choice Questions (MCQ) directed to Case A (without learning interventions) and Cases B-C (with learning interventions)

Groups		Learrning Interventions			PreTest		PostTest 1 MCQ Case A		PostTest 2 MCQ Case A		PostTest 3 MCQ Case A		PostTest 1 MCQ Case B-C		PostTest 2 MCQ Case B-C		PostTest 3 MCQ Case B-C	
	n	VS	IA	GQ	Mean %	SD	Mean %	SD	Mean %	SD	Mean %	SD	Mean %	SD	Mean %	SD	Mean %	SD
1	23				76.5	17.5	95.7	11.5	69.6	26.4	51.5	27.7	76.1 ***	14.6	73.9	19.6	52.7	16.4
2	22			y	71.8	22.8	83.3	22.4	72.7	24.4	50.0	25.6	66.5 **	66.5	60.2 *	16.2	46.0	12.4
3	22		y		73.2	15.5	81.8	24.6	74.2	22.8	48.9	25.0	69.3	13.8	63.1	20.9	47.7	15.8
4	21		y	y	78.6	16.8	90.5	21.5	79.4	22.3	52.4	28.4	67.9 ***	17.5	64.3 *	18.7	45.8	18.7
5	21	y			79.1	17.9	88.9	19.2	81.0	19.9	44.1	22.2	74.4 *	15.5	70.8	22.1	56.0	19.6
6	20	y		y	75.5	17.3	78.3	27.1	71.7	24.8	48.8	26.3	70.6	12.4	70.0	15.4	48.8	48.8
7	22	y	y		73.6	19.9	84.9	17.0	78.8	19.4	39.8	29.5	64.8 ***	20.3	68.2	18.4	51.7	51.7
8	22	y	y	y	77.7	16.6	86.4	19.7	81.8	19.9	48.9	30.4	65.9 **	15.0	67.1 *	22.3	49.4	20.6
Total	173				75.7	18.0	86.3	20.0	76.1	22.6	48.0	26.7	69.4	15.7	67.2	19.5	49.8	17.7

Note. VS (virtual slides), IA (interactive annotations, GQ (guiding questions).

One-way ANOVA between groups, $p > .05$ for PreTest and PostTests 1-3 Case A and Cases B-C.

Paired samples t-tests within groups compares scores on MCQ directed to Case A and Cases B-C in PostTests 1-3, * $p < .05$, ** $p < .005$, *** $p < .001$.

PostTest 1-3 Case A

The results of between group PostTest scores on MCQ directed to Case A are shown according to student CAI module group assignment in Table 5.

The mean scores for all 173 students on the questions directed at Case A in the PostTests after the lab exercise were 86.3 ± 20.9% on PostTest1, 76.1 ± 22.6% on PostTest2 and 48.0 ± 26.7% on PostTest3. Results were analyzed between groups by one-way ANOVA. There were no statistically significant differences between the different groups for any of the PostTest MCQ directed to Case A, one-way Welch's ANOVA PostTest 1 $F(7,165) = 1.485$, $p = .176$, $\eta^2 = .059$; PostTest 2 $F(7,165) = .919$, $p = .493$, $\eta^2 = .037$; and, PostTest 3 $F(7,165) = .503$, $p = .831$, $\eta^2 = .021$.

A Kruskal-Wallis H test was also run in light of some violations of assumptions for ANOVA testing to determine if there were any significant differences in test results between the groups in the PostTest 1-3 scores with MCQ directed to Case A. Median test scores were not statistically significantly different between groups on any of the tests, PostTest 1 result X2(7) = 10.835, $p = .146$, PostTest 2 result X2(7) = 5.394, $p = .612$, PostTest 3 result X2(7) = 3.658, $p = .467$.

The results indicate that student group assignment did not have a significant effect on the test results for questions directed at Case A at any time point after the CAI lab exercise, indicating that student groups were equivalent in performance on test questions related to Case A, which was presented in a similar manner without interventions in all of the CAI modules.

PostTest 1-3 Cases B-C

The learning intervention variables assigned to each of the student groups and the results of PostTest scores on Cases B and C-directed questions according to student CAI module group assignment are shown in Table 5.

The mean scores for all 173 students on the questions directed at Case B and C in the PostTests after the lab exercise were 69.4 ± 15.7% on PostTest 1, 67.2 ± 19.5% on PostTest 2 and 49.8 ± 17.7% on PostTest 3. Results were analyzed between groups by one-way ANOVA. One-way ANOVA showed no statistically significant difference between the different groups for any of the PostTest MCQ directed at Cases B-C: PostTest 1 $F(7,165) = 1.519$, $p = .164$, $\eta^2 = .061$; PostTest 2 $F(7,165) = 1.185$, $p = .314$, $\eta^2 = .048$; and, PostTest 3 $F(7,165) = .836$, $p = .559$, $\eta^2 = .034$.

A Kruskal-Wallis H test was also run in light of some violations of assumptions for ANOVA testing to determine if there were differences in test results scores between the groups in the PostTest 1-3 scores with MCQ directed to Cases B=C. Median test scores were not statistically significantly different between groups on any of the tests, PostTest 1 result X2(7) = 9.112, $p = .245$, PostTest 2 result X2(7) = 9.102, $p = .245$, PostTest 3 result X2(7) = 6.642, $p = .467$.

In summary, these results indicate that student group assignment did not have a significant effect on the test results for questions directed at Cases B and C at any time point after the CAI lab exercise. The results indicate that none of the interventions or their combination in the individual CAI learning modules had an impact that could be measured by exam performance of the student groups. These results indicate that the learning interventions presented in the CAI modules in Cases B and C did not improve the performance of students on PostTest questions directed to those cases. It is noted, however, that the results of PostTest scores for students in Group 1, who had no interventions expressed in their assigned CAI module, tended to be higher for PostTests 1 and 2 than for students in Groups 2-8, who had interventions included in their assigned CAI modules.

PostTest 1-3 Case A vs. Cases B-C

These results compared student test performance within groups on MCQ directed to Case A or Cases B-C in PostTests 1-3. Although results on tests between student groups had shown no statistically significant differences between the student groups assigned to the eight CAI modules, student performance within groups did show consistently higher scores on questions directed to Case A as compared to questions directed to Cases B-C (Table 5). A Paired-Samples t-test was used to determine statistically significant differences within groups for student group scores on PostTests 1-3, comparing multiple choice questions (MCQ) directed to Case A (without learning interventions) and MCQ directed to Cases B-C (with learning interventions). Results of paired sample differences were greatest for PostTest 1. Questions directed to Case A elicited statistically significant higher scores than questions directed to Cases B-C in all groups 1-8, except groups 3 and 6 (Table 5). The greatest significant differences between scores were seen in groups 1, 4 and 7 as follows. In Group 1, test results of MCQ for Case A (M = 95.7%, SD = 11.5%) and MCQ for Cases B and C (M = 76.1%, SD = 14.6%) showed a statistically significant difference of 19.6%, CI [11.2, 27.9], $t(22) = 4.868$, $p < .001$, $d = 1.02$. In group 4, test results of MCQ for Case A (M = 90.5%, SD = 21.5%) and MCQ for Cases B-C (M = 67.9%, SD = 17.5%) showed a statistically significant difference of 22.6%, CI [15.5, 29.7], t(20) = 6.659, p < .001, d = 1.45. In group 7, test results of MCQ for Case A (M = 84.9.5%, SD = 17.0%) and MCQ for Cases B and C (M = 64.8%, SD = 20.3%), showed a statistically significant difference of 20.1%, CI [11.8, 28.4], t(21)= 5.034, p <.001, d = 1.07.

In summary, these results comparing within group PostTest scores showed statistically significant higher scores in PostTest 1 on questions directed to Case A, which were presented in all CAI modules with no learning interventions, than for questions directed to Cases B-C which were presented in CAI modules with various learning interventions, although it did not reach statistical significance for Groups 3 and 6.

PostTest 1-3 Sequential Scores

Exams after the CAI lab exercise were administered immediately after the lab exercise (PostTest 1), two weeks after the lab exercise (PostTest 2), and two months after the lab exercise (PostTest 3). The results of PostTest scores on MCQ directed to Case A (without learning interventions) and to Cases B-C (with learning interventions) are shown in association with student CAI learning module group assignment (Groups 1-8) in Tables 5-7 and Figures 2 and 3.

Student performance for all student groups decreased progressively over time, as measured by mean scores on successive PostTest exams 1-3, with MCQ directed at Case A (without learning interventions) or Cases B-C (with learning interventions).

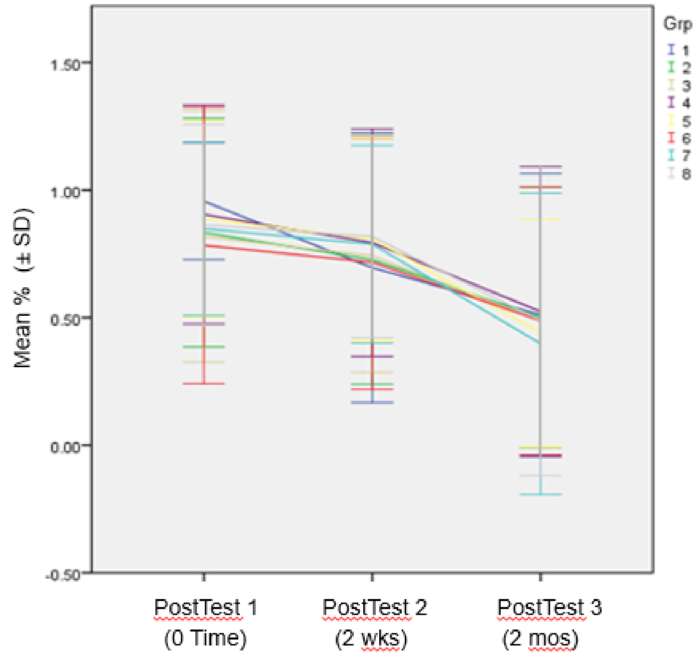


Figure 2. Mean % ± SD Student Group Scores on Multiple Choice Questions Directed to Case A (without learning interventions) in PostTests 1-3 after The Computer-Assisted Instruction Laboratory Exercise

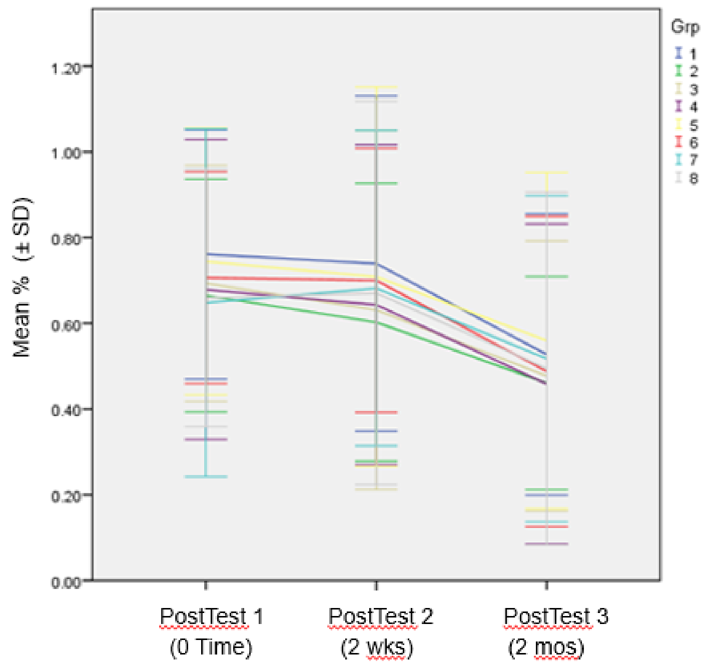


Figure 3. Mean % ± SD Student Group Scores on Multiple Choice Questions Directed to Cases B-C (with learning interventions) in PostTests 1-3 after the Computer-Assisted Instruction Laboratory Exercise

A one-way Repeated Measures ANOVA was conducted with pairwise comparisons between PostTest scores over time on MCQ directed to Case A for all students in the class (n= 173). There were statistically significant changes in the PostTest scores over time, $F(1.932, 332.348) = 135.419$, $p < .001$, partial $\eta^2 .44$, with scores decreasing from 86.3% ± 20.0 on PostTest 1 to 76.1% ± 22.6 on PostTest 2 and to 48.0% ± 26.7 on PostTest 3. Post hoc analysis with a Bonferroni adjustment revealed that scores on PostTest 3 with MCQ directed to Case A were statistically significantly decreased from PostTest 1 by - 38.3% (95% CI, -44.4 to -32.2) , $p < .001$), and significantly decreased from PostTest 2 by -28.1% (95% CI, -34.2 to -22.0) , $p < .001$), and PostTest 2 had a significant decrease from PostTest 1 of -10.2% (95% CI, -15.5 to -5.0), $p = <.001$.

The results comparing student performance on applied knowledge using CAI modules with or without embedded learning interventions (Cases B-C) also showed a significant decrease in performance of all students over time, using one-way Repeated Measures ANOVA with pairwise comparisons of PostTest results on MCQ directed to Cases B-C (with learning interventions) over time for all students in the class (n= 173). There were statistically significant changes in the PostTest scores over time, (2, 344) = 85.653, $p < .001$, partial $\eta^2 .33$, with scores decreasing from 69.4% ± 15.7 on PostTest 1 to 67.2% ± 19.5 on PostTest 2 and to 49.8% ± 17.1 on PostTest 3. Post hoc analysis with a Bonferroni adjustment revealed that scores on PostTest 3 for MCQ directed to Cases B-C were statistically significantly decreased from PostTest 1 by -19.7% (95% CI, -23.6 to -15.8), $p < .001$), and significantly decreased from PostTest 2 by -17.4% (95% CI, -21.4

to -13.4) %, $p < .001$), but there was not a significant decrease between PostTest 2 and PostTest 1 at -2.2% (95% CI, -6.2 to -1.8), $p = .530$).

The within group results comparing performance on MCQ directed specifically to Case A or to Cases B-C using Pairwise Comparisons by t-test showed no significant differences in mean scores within groups when comparing results for PostTest 1 and PostTest 2 on questions directed to Case A or on questions directed to Cases B and C ($p < .05$), except group 1, which showed a significant difference in scores on questions directed to Case A (($p < .05$) (see Table 6 and 7).

Table 6

Mean % ± SD Within Student Group Scores on PostTest 1-3 Multiple Choice Questions (MCQ) directed to Case A (without learning interventions)

Groups		Learrning Interventions			PostTest 1 MCQ Case A		PostTest 2 MCQ Case A		PostTest 3 MCQ Case A	
	n	VS	IA	GQ	Mean %	SD	Mean %	SD	Mean %	SD
1	23				95.7 * ° ^	11.5	69.6 * °	26.4	51.5 * ^	27.7
2	22			y	83.3 * ^	22.4	72.7 ° ^	24.4	50.0 * ° ^	25.6
3	22		y		81.8 * ^	24.6	74.2 ° ^	22.8	48.9 * ° ^	25.0
4	21		y	y	90.5 * ° ^	21.5	79.4 ° ^	22.3	52.4 * ° ^	28.4
5	21	y			88.9 * ^	19.2	81.0 ° ^	19.9	44.1 * ° ^	22.2
6	20	y		y	78.3 * ^	27.1	71.7 ° ^	24.8	48.8 * ° ^	26.3
7	22	y	y		84.9 * ^	17.0	78.8 ° ^	19.4	39.8 * ° ^	29.5
8	22	y	y	y	86.4 * ^	19.7	81.8 °^	19.9	48.9 * ° ^	30.4
Total	173				86.3	20.0	76.1	22.6	48.0	26.7

Note. VS (virtual slides), IA (interactive annotations, GQ (guiding questions).
Repeated measures ANOVA within groups and pairwise comparisons by t-test on MCQ directed to Case A
* $p < .05$, ° $p < .05$, ^ $p < .05$

It is noted that Group 1 had a higher mean score on PostTest 1 MCQ directed to Case A than any other group (96% compared to 78-90%), which may explain the results showing a significant decrease in student scores between Postest 1 and PostTest 2. Group 1 also showed no significant difference in mean scores on questions directed to Case A between results on PostTest 2 and PostTest 3, whereas every other group did show a significant difference on questions directed to Case A and also on questions directed to Cases B-C (p< .05).

Table 7

Mean % ± SD Within Student Group Scores on PostTest 1-3 Multiple Choice Questions (MCQ) directed to Cases B-C (with learning interventions)

Groups	n	Learning Interventions VS	IA	GQ	PostTest 1 MCQ Case B-C Mean %	SD	PostTest 2 MCQ Case B-C Mean %	SD	PostTest 3 MCQ Case B-C Mean %	SD
1	23				76.1 * ^	14.6	73.9 ° ^	19.6	52.7 * ° ^	16.4
2	22			y	66.5 * ^	66.5	60.2 ° ^	16.2	46.0 * ° ^	12.4
3	22		y		69.3 * ^	13.8	63.1 ° ^	20.9	47.7 * ° ^	15.8
4	21		y	y	67.9 * ^	17.5	64.3 ° ^	18.7	45.8 * ° ^	18.7
5	21	y			74.4 * ^	15.5	70.8 ° ^	22.1	56.0 * ° ^	19.6
6	20	y		y	70.6 * ^	12.4	70.0 ° ^	15.4	48.8 * ° ^	48.8
7	22	y	y		64.8 * ^	20.3	68.2 ° ^	18.4	51.7 * ° ^	51.7
8	22	y	y	y	65.9 * ^	15.0	67.1 ° ^	22.3	49.4 * ° ^	20.6
Total	173				69.4	15.7	67.2	19.5	49.8	17.7

Note. VS (virtual slides), IA (interactive annotations, GQ (guiding questions).
Repeated measures ANOVA within groups and pairwise comparisons by t-test on MCQ directed to Cases B-C
* $p < .05$, ° $p < .05$, ^ $p < .05$

Effect of Student Class Rank on Exam Performance

When results emerged that did not show significant effects of the learning interventions on student group exam performance, class rank was added as an additional variable for study with the idea that the learning interventions might have a differential effect on students with a high or low aptitude for leaning pathology.

Class Rank by Quintile Score

In order to determine the effect of class rank, students were grouped into one of three quintile ranks based on their cumulative average test scores (Cum Ave) at the time of PostTest 2, Low Quintile, Mid-Quintiles and High Quintile. When students were sorted into Low-, Mid-, and High-Quintile ranks on the basis of their Cum Ave, significant differences were found in their scores on the PreTest and PostTests 1-3 exams with MCQ directed to Case A (without learning interventions) or to Cases B-C (with learning interventions), as determined by One-Way ANOVA (see Table 8).

Welch's one-way ANOVA showed statistically significant differences between the high-, mid- and low-quintile groups determined by class rank for all PostTest Cases with MCQ directed to Case A (without learning interventions) and Cases B-C (with learning interventions) (p <.001), except for PostTest 3 MCQ directed to Case A (p= .106). Games-Howell post-hoc analysis revealed statistically significant differences between High and Low Quintile class rank for PostTest 1 and 2 MCQ directed to Case A and to Cases B-C and Middle Quintiles to Low Quintile class rank for MCQ directed to Cases B and C on PostTests 1, 2 and 3. Significant differences and confidence intervals are included in more detail in Appendix D.

Class rank as a Continuous Score

A Pearson correlation coefficient was calculated to determine relationships between class rank at the time of the CAI lab exercise (Cum Ave) and scores achieved on the PreTest and PostTests 1-3. Class rank was analyzed as a continuous variable. A significant positive correlation was found in every case, moderate for PreTest, $r_s(171) = .467$, p < .001; moderate to strong for PostTests 2 and 3 with questions directed at Cases B-C (with learning interventions), $r_s(171) = .644$, $p < .001$ for PostTest 2 and $r_s(171) = .407$, $p < .001$ for PostTest 3; moderate to weak for PostTest 1 and 2 with questions directed to Case A (without learning interventions), $r_s\ (171) = .269$, $p < .001$ for PostTest 1 and $r_s\ (171) = .181$, $p < .001$ for PostTest 2 and for PostTest 1 with questions directed to Cases B-C, $r_s\ (171) = .290$, $p < .001$.

In summary, these results indicate that there is a correlation between student overall performance in Pathology and the scores that students achieve on PreTest and PostTests 1-3 associated with the CAI lab exercise.

Table 8

Mean % ± SD Student Performance Scores Between-Quintiles of Class Rank , as Determined by Cum Ave at CAI, on the PreTest , on PostTest 1-3 Multiple Choice Questions (MCQ) Directed to Case A (without learning interventions), on PostTest 1-3 MCQs Directed to Cases B-C (with learning interventions)

Student Group	n	Cum Ave Mean % ± SD	PreTest Mean % ± SD	PostTest 1 MCQ Case A Mean % ± SD	PostTest 2 MCQ Case A Mean % ± SD	PostTest 3 MCQ Case A Mean % ± SD	PostTest 1 MCQ Case B-C Mean % ± SD	PostTest 2 MCQ Case B-C Mean % ± SD	PostTest 3 MCQ Case B-C Mean % ± SD
Hi Quintile	35	90.4 (2.4) *	88.6 (10.3)*	94.3 (15.1) * ~'	78.1 (11.9)	67.1 (23.3) * ~°	77.1 (11.9) *	84.6 (14.6) *	61.4 (17.8) *
Mid-Quintile	105	80.8 (4.6) *	75.0 (17.6)*	85.4 (10.3) ~°	78.1 (22.6)	44.0 (25.6) ~°	69.9 (15.1) *	67.1 (16.7) *	48.4 (16.5) *
Lo Quintile	33	70.0 (4.0) *	64.2 (17.1)*	80.8 (.21) *	67.7 (25.7)	40.1 (25.0) *	59.8 (16.5) *	48.9 (14.8) *	40.5 (14.7) *
Total	173	80.7 (7.6)	75.7 ± 18.0	86.3 ± 20.0	76.1 ± 22.6	48.0 ± 26.7	69.4 ± 15.7	67.2 ± 19.5	49.8 ± 17.7

Note. Cum Ave scores is class mean score at the time of PostTest 2 and is used for determinations of class rank.

* $p < 0.05$, ° $p < 0.05$, ~ $p < 0.05$ for between Quintile comparisons of all tests.

Predictive Ability of Learning Interventions and Class Rank

Multiple regression analysis was done to determine if the learning interventions - Virtual Slides, Interactive Annotations and Guiding Questions, could predict student scores on MCQ directed to Cases B-C in PostTests 1-3. Initial results showed no significant positive predictive ability for any of the three learning interventions, and a slight negative effects of interactive annotations (data not shown). In light of the other results that had shown positive correlations of student class rank with exam performance, class rank was added as a variable to the multiple regression analyses to determine if it had predictive ability for student exam performance. In order to include class rank as a variable or variables in multiple regression analyses, class rank was considered as two separate variables of High-Quintile and Low-Quintile class rank and considered in other analyses as a continuous variable. In addition, the predictive ability of the three learning interventions on PostTest 1 exam scores for MCQ directed to Cases B-C was tested within High-Quintile and Low-Quintile student groups. The results of these studies are described below, with additional details in Appendix D.

Learning Interventions and Class Rank by Quintile Score

Multiple regression analysis of PostTest 1 found that Quintile (High or Low), Virtual Slides, Interactive Annotations and Guiding Questions could statistically significantly predict scores on questions directed to Cases B-C in PostTest 1, $F(5, 167) = 6.450$, $p < .001$, adj. $R^2 = 0.137$. The variables that were statistically significant predictors included: High Quintile Cum Ave as a positive predictor (unstandardized correlation coefficient -.073, $p = .012$); Low Quintile Cum Ave as a negative predictor (unstandardized correlation coefficient - .101, p= .001); and Interactive Annotations as a negative predictor (unstandardized correlation coefficient - .051, *p*= .023).

Multiple regression analysis of PostTest 2 found that the five variables that were entered could statistically significantly predict scores on questions directed to Cases B-C in PostTest 2, $F(5, 167) = 18.326$, $p < .001$, adj. $R^2 = .335$. The only variables that were statistically significant predictors included: High Quintile Cum Ave as a positive predictor (unstandardized correlation coefficient 0.173, $p < .001$) and Low Quintile as a negative predictor (unstandardized coefficient - 0.182, $p < .001$).

Multiple regression analysis of PostTest 3 found that the five variables that were entered could statistically significantly predict scores on questions directed to Cases B-C in PostTest 3, $F(5, 167) = 6.864$, $p < .001$, adj. $R^2 = .146$. The only variables that were statistically significant predictors included: High Quintile Cum Ave as a positive predictor (unstandardized correlation coefficient .123, $p < .001$) and Low Quintile as a negative predictor (unstandardized coefficient - .083, $p = .012$).

Learning Interventions and Class Rank as Continuous Score

When student rank was entered as a continuous variable in analysis of variables affecting PostTest 1, the analysis found $F(4,168) = 6.088$, p < .001, adj. $R^2 = .106$ and the unstandardized correlation coefficient was .608, p < .001). When Class Rank was entered as a continuous variable in analysis of variables affecting PostTest 2 and 3, there was no statistically significant prediction of scores on questions directed to Cases B-C, PostTest 2 $F(4,168)=1.113$, p= .352, adj. $R^2 = .003$. PostTest 3 $F(4,168)=1.594$, p= .178, adj. $R^2 =$ .014.

Predictive ability of Learning Interventions Within Quintile Class Ranks

In order to determine if the three learning interventions might have a disproportionate effect on students with a High, Mid-level or Low Quintile Cum Ave, multiple regression was also applied specifically to PostTest 1 scores of students within these three groups. Multiple regression for the three variables in students with a High Quintile Class Rank (n= 35) was not statistically significant $F(3, 31) = .309$, $p = .819$, nor was it significant for students with Mid-Quintiles Class Rank (n= 105) $F(3, 101) = 1.055$, $p = .372$. Multiple regression for the three variables in students with a Low Quintile Cum Ave (n= 33) was statistically significant $F(3, 29)= 3.165$, p= .039, but the correlation was a negative one for Interactive Annotations (unstandardized correlation coefficient - .146, p= .014).

In summary, these results indicate that student rank as determined by Cum Ave at the time of the Computer-Assisted Instruction(CAI) lab exercise and analyzed as a continuous variable was a strong predictor of performance on PostTest 1, but not on PostTests 2 and 3. Student rank as determined by a High Quintile Cum Ave was a weak positive predictor of performance on all Post-Tests 1-3 and a Low Quintile Cum Ave was a weak negative predictor. None of the learning interventions were associated with any positive predictive value, but Interactive Annotations was a very weak negative predictor of performance for students with a Low Quintile Cum Ave.

Student Survey

The survey of student learning and satisfaction included ten questions to be answered on a Likert-scale and three open-ended questions. A summary of the numbers of students responding to each Likert-style question is provided in Table 9 and the numbers responding to the open-ended questions is summarized in Appendix C.

Likert-style Survey Questions

The anonymous student survey concerning the CAI modules and the lab exercise included ten questions answered on a six-point Likert-type scale (1 strongly agree to 6 strongly disagree) and four open-ended questions for comment. Surveys were completed immediately after the laboratory exercise.

Results showed no significant differences of Likert-type scale scores between groups on any question, as assessed by Kruskal-Wallis nonparametric testing for 143 students who answered these survey questions (Appendix C). Median test scores were not statistically significantly different between groups on any of the questions, Question 1 result X2(7) = 5.173, p = .639, Question 2 result X2(7) = 5.231, p = .632, Question 3 result X2(7) = 7.656, p = .364, Question 4 result X2(7) = 6.951, p = 434, Question 5 result X2(7) = 10.089, p = .184, Question 6 result X2(7) = 2.058, p = .957, Question 7 result X2(7) = 2.556, p = .923, Question 8 result X2(7) = 6.114, p = .526, Question 9 result X2(7) = 9.700, p = .206, Question 10 result X2(7) = 10.143, p = .181. Means and medians of the Likert scale scores between student groups are presented in Table 9. The scale had a high level of consistency for all students, as determined by a Chronbach's alpha = .981.

Table 9

Means, SD and Medians of Student Group Responses to Ten Survey Questions on Use of a Computer-Assisted Instruction Learning Module

Student Groups	n		Likert Q1	Likert Q2	Likert Q3	Likert Q4	Likert Q5	Likert Q6	Likert Q7	Likert Q8	Likert Q9	Likert Q10
1	22	Mean	1.82	2.23	1.82	2.09	2.18	2.23	2.23	2.45	2.36	2.59
		SD	0.80	1.19	0.80	1.11	1.53	1.38	1.57	1.22	1.59	1.65
		Median	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
2	21	Mean	1.95	2.14	2.00	2.19	2.43	2.10	2.00	2.52	2.33	2.38
		SD	0.86	1.01	0.95	1.21	1.25	1.04	0.95	1.47	1.15	1.20
		Median	2.00	2.00	2.00	2.00	2.00	2.00	2.00	3.00	3.00	3.00
3	19	Mean	1.74	2.68	2.32	2.95	3.11	2.74	2.58	2.95	3.05	3.16
		SD	1.41	1.38	1.20	1.68	1.94	1.66	1.57	1.68	2.09	1.80
		Median	1.00	2.00	2.00	3.00	3.00	3.00	2.00	3.00	3.00	3.00
4	17	Mean	1.76	2.29	2.24	2.41	2.65	2.29	2.35	2.76	2.88	2.29
		SD	0.90	1.16	1.15	1.12	1.50	1.26	1.17	1.35	1.41	1.21
		Median	2.00	2.00	2.00	2.00	2.00	2.00	2.00	3.00	3.00	2.00
5	21	Mean	1.86	2.43	2.29	2.67	3.00	2.19	2.14	2.71	3.00	3.19
		SD	0.57	0.93	0.90	1.28	1.67	0.93	1.01	1.45	2.05	1.86
		Median	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	3.00
6	12	Mean	1.75	2.50	2.00	2.83	3.75	2.08	2.25	3.50	3.92	3.58
		SD	0.75	1.51	0.95	1.03	1.66	0.90	0.87	1.31	1.31	1.78
		Median	2.00	2.50	2.00	3.00	4.00	2.00	2.00	3.00	3.50	3.00
7	16	Mean	1.75	1.88	1.63	2.25	2.44	2.06	2.00	2.56	2.44	2.06
		SD	0.86	0.81	0.72	1.13	1.26	1.00	0.82	1.59	1.41	1.53
		Median	2.00	2.00	1.50	2.00	2.00	2.00	2.00	2.00	2.00	1.50
8	15	Mean	1.93	2.07	1.87	2.40	2.53	2.13	2.07	2.40	2.80	2.67
		SD	0.59	1.03	0.99	1.12	1.36	1.13	1.03	1.45	1.74	1.68
		Median	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	3.00
Total	143	Mean	1.83	2.28	2.03	2.45	2.71	2.24	2.20	2.70	2.79	2.72

Note. Likert Q are questions with scale scores ranging 1-6 (1 - strongly agree, 2 - moderately agree, 3 - slightly agree, 4 - slightly disagree, 5 - moderately disagree, 6 - strongly disagree)

No significant differences in results between student groups, p >.05

Chronbach's aplha = .981

Results of the survey scores that were 1-2 on the Likert-type scale are as follows; all scores of 5-6 are less than 10%, except as noted.

On questions about perceived learning using the Computer-Assisted Instruction (CAI) modules (n = 143 students),

(1) 81.0% of students thought purpose and objectives were clear

(2) 72.8% thought the information was presented clearly

(3) 65.0% thought the CAI modules were correctly set to the student's level of learning

(4) 66.6% thought the modules complemented other learning materials in the Pathology course.

(5) 45.1% thought they learned better from the CAI modules than the traditional in-class computer lab and small group; 17.6% scored 5-6 on the Likert-type scale

(6) 50.2% thought that they learned better from the CAI modules than from traditional study in-class with virtual slides and PPT cases; 17.3% scored 5-6 on the Likert-type scale

On questions of satisfaction with use of the CAI modules:

(7) 48.9% enjoyed learning from the CAI modules; 11.4 % scored 5-6 on the Likert-type scale

(8) 51.8% would like to use similar CAI modules to learn other pathology; 14.5% scored 5-6 on the Likert-type scale.

(9) 48.5% would recommend the CAI modules to their peers

On the question concerning the CAI module technology

(10) 66.2% thought the CAI module was easy to navigate

Open-Ended Survey Questions

Students entered responses to four open-ended questions on the survey concerning the Computer-Assisted Instruction (CAI) modules and lab exercise (n= 144). The numbers of students responding to each question in each group are presented in Appendix C. The questions for comment were as follows:

(1) What was very good/helpful about the Computer-Assisted Instruction (CAI)?

(2) What was very bad/distracting about the CAI?

(3) What suggestions do you have to improve the CAI?

Positive and negative themes that came out of student comments were not elicted by direct questions concerning any of the learning interventions, since not all students were exposed to each intervention. All student comments about features of the CAI learning modules were anonymous and spontaneous.

Positive themes

The case-based nature of the exercise was identified as a positive feature of the CAI learning module by 20 students out of 115 who responded (17%), 1-6 students in each Group.

- "This was very well done. It gave me clinical context, let me think on my own, and then presented information in a clear, succinct manner that allowed for easy learning." (from a student in Group 6)

- “These were useful to put the pathology and clinical presentations into perspective, and show how the knowledge is used clinically.” (from a student in Group 7)
- “Having the pathology relate back to a case made it more relevant and enhanced its clinical relvance.” (from a student in Group 3)

2. Aggregation of information in one place was identified as a positive feature of the CAI learning module by 29 students out of 115 who responded (25%), 1-6 students in each Group.
   - “It is nice to have all of the information that we will be responsible for in one place.” (from a student in Group 1)
   - “very informative in the sense that it gave me all the information I would need to know about for the case and that it brought it slides from lectures so that it became an active learning rather than passive.” (from a student in Group 2)
   - “I thought it was helpful to have the case presented and then everything explained (from the physical findings to the lab results). These were very effective” (from a student in Group 5)
3. Self-paced independent study was identified as a positive feature of the CAI learning module by 17 students out of 115 who responded (15%), 2-4 students in 7 of the 8 Groups.
   - “I like that I was able to go at my own pace” (from a student in Group 1)
   - “Could be used at one's own pace and within one's own schedule” (from a student in Group 8)
   - “It was nice to be able to work at you own pace” (from a student in Group 6)

4. Slide labels (annotations) were identified as a positive feature of the CAI learning module by 20 students out of 115 who responded (17%), 2-6 students in 7 of the 8 Groups. No comments specifically commented on the interactivity of the slide labels.
   - "All the pathology was clearly labeled, which made it easy to recognize." (from a student in Group 1)
   - "I really liked how you had micro images and areas highlighted with arrows pointing out key features." (from a student in Group 2)
   - "pathology slides with specific pathology labeled so I don't have to go looking for it online or in a text book" (from a student in Group 2)
   - "I liked the fact that the virtual slides were clear and labeled" (from a student in Group 7)
5. Questions interactive with answers were identified as a positive feature of the CAI learning module by 10 students out of 54 who had CAI with this feature and who responded (18%), 1-3 students in each of the four Groups with interactive questions made positive comments.
   - " I liked how it presented cases and asked questions then gave answers with some background information if pertinent." (from a student in Group 2)
   - "I loved how there was a question and then an answer given." (from a student in Group 4)
   - " It gave me a chance to answer the questions while also having the answers handy so that I could learn. Having the information easily available made studying much easier. I would be very excited to keep learning this way." (from a student in Group 4)

- " Asking questions about the findings (e.g. Why is AST > ALT in ASH ?) before providing the answers promoted active learning." (from a student in Group 8)

6. Hyperlinks were identified as a positive feature of the CAI learning module by 16 students out of 115 who responded (14%), 1-3 students in each Group.
    - " hyperlinked powerpoints helped to integrate" (from a student in Group 1)
    - "I really liked the background hyperlinks because even when I remembered the information it was helpful to review it in context." (from a student in Group 2)
    - "I liked the extra links that showed basic pathology and information that we should learn to understand the case" (from a student in Group 5)
7. Other positive comments: There were many additional positive comments on the thoroughness of information provided in the CAI or the systematic nature of presentation. Two students (Group 1) made direct comments about the ability of the CAI to standardize the learning materials for all students.

Negative themes

Technical difficulties were identified as a negative feature of the CAI learning modules by 49 students out of 117 who responded (42%), 1-3 students in every Group. Nearly all of the comments concerned the ability to seamlessly access slides with additional information available as hyperlinks when using their own computers.

- "The tech issues were very frustrating, and took away from the learning experience. If these issues could be worked out, I think this method of learning pathology could be very effective." (from a student in Group 1)

- “Hyperlinks not working properly, difficult to use annotated slides. Overall, IT aspect of the CAI was not satisfactory, taking away from learning.” (from a student in Group 2)
- “Constantly having to open and close powerpoint windows because hyperlinks were not functioning properly. At times I would find myself with 6 powerpoint/internet explorer tabs open which made it slightly overwhelming to progress linearly through the module” (from a student in Group 6)

2. Density of material in allotted time was identified as a negative feature of the CAI learning module by 21 students out of 117 who responded (18%), 2-7 students in every Group.
   - “There was a lot of information on every slide, so it took me almost the entire two hours to get through all three cases.” (from a student in Group 1)
   - “It was a lot of information to cover in 2.5 hours” (from a student in Group 4)
   - “Information was too dense especially when we are going to be tested on the material at the end of the session.” (from a student in Group 8)
3. Classroom setting or length or time required in class were identified as a negative feature of the CAI learning module by 34 students out of 117 who responded (29%), 2-8 students in every Group made negative comments.
   - “Being in class was distracting. It should be something accessible from home that we can do in our own time.” (from a student in Group 3)
   - “would be nice to have it to do at home or in the library for self-review rather than in a classroom setting.” (from a student in Group 4)

- "Very difficult to stay focused and read in a room of 150+ people, Need to be focused/in the right mood to do that much reading, go through all those cases efficiently, Can't do this in a large group setting, should be done at home" (from a student in Group 7)
- "The total exercise was FAR too long for my attention span at a computer"

4. Social isolation was identified as a negative feature of the CAI learning module by 9 students out of 117 who responded (8%), 1-3 students in 6 of the 8 Groups.
   - "I didn't like not being able to discuss things with my peers." (from a student in Group 5)
   - "I thought it was difficult to not be able to talk to anybody. I couldn't ask questions or discuss anything to help clarify" (from a student in Group 6)
   - "For me it was too long and too quiet. I like to be able to discuss the cases." (from a student in Group 7)
5. Other negative comments:

   Two students (Group 3, without interactive questions) suggested that interactive questions would have been helpful. Four students (one in Group 2 and 3 in Group 4) felt that they needed expert guidance. One student (Group 5) felt that an auditory component was needed.

Summary of Results

In summary, the results do not support hypotheses 1-3 which predicted that the learning interventions would have a positive effect on student performance. No learning intervention had a significant positive effect on student exam performance and, in fact, interactive image annotations had a weak negative predictive effect by multiple regression in Post Test 1. Student class rank showed the strongest correlation with performance on PostTest 1. The learning interventions also showed no positive effect on retention of information, since there was a significant decrease in student scores on successive exams (Post Tests 2 and 3) in all groups. The hypothesis that students would perceive the interventions helpful to learning was supported in part by results of the survey. The use of guiding questions was perceived to have a positive impact on learning. Slide label annotations were also perceived to have a positive impact on learning, but the interactive quality of the annotations was not specifically noted. The intervention of virtual slides, rather than a static image, was also without comment.

The hypothesis that student satisfaction and perceived learning would show a positive correlation with features that gave students more control over their learning within the CAI module was supported. The opinions reflected in the Likert-type scale survey suggest that approximately half of the students felt that self-study using the CAI learning modules was complementary to other learning materials on the subject and that they were enjoyable and preferable to the usual in-class lab exercise and small groups. Positive themes in the survey concerned satisfaction with the ability to control the pace of learning, interactivity with the program, use of annotations, guiding questions and self-selection of informational content through hyperlinks. Negative comments by students concerning use of the the CAI learning modules centered around technical issues that some students experienced in the smooth operation of the program on their personal computers and the stress of using the CAI learning module for self-study in a time-limited, class room setting with the pressure of a test immediately on completion of the exercise.

CHAPTER 5: DISCUSSION

The experimental study reported here tested the impact on student learning of three instructional design interventions in the configuration of a case-based, self-study, Computer-Assisted Instruction (CAI) learning module used as an enhancement in a Year-2 medical school Pathology course laboratory exercise. Interventions included the use of virtual slides (VS) instead of static images; the use of interactive image annotations (IA) instead of static labels; and, the use of guiding questions (GC) before providing essential information. Each of the three interventions was controlled for by assigning students to one of eight CAI learning modules in which one case (A) was in common with no interventions and two cases (B and C) was presented with one or more of the interventions to be tested. The impact of the CAI exercise was determined by student performance on MCQ tests given immediately after the exercise (Post Test 1), 2 weeks after the exercise (Post Test 2) and 2 months after the exercise (Post Test 3). Student opinion on perceived learning and satisfaction in the use of the CAI program was also assessed with a series of Likert-type scale and open-ended questions.

Discussion of Findings

The study questions and hypotheses predicted that the learning interventions would have a positive impact on student learning, as measured by exam performance, and that students would perceive that their learning was enhanced by use of the learning interventions in the CAI modules.

The results show that the learning interventions did not improve student learning, as measured by exam performance, and did not improve retention of learning over time, as measured in successive exams (Post Tests 1-3). In fact, the interactive annotations showed a weak negative correlation with student performance on Post Test 1 when the interventions were analyzed as variables in multiple regression, especially for students in the lowest quintile of class rank at the time of the exam.

The idea that the interventions may have added too much complexity for a timed, in-class CAI lab exercise is supported by the observation that all eight student groups using all of the variously configured CAI modules showed significantly better scores on questions directed to Case A, which was presented with no interventions in all CAI modules, than those questions directed to Cases B and C, which were presented with various combinations of the learning interventions. It is possible, however, that students found Case A to be easier to learn from or that the questions for Case A were easier to answer. This idea is supported by the finding that students in Group 1 also had significantly higher scores on questions directed to Case A than for those directed to Cases B and C, even though Cases B and C were presented without interventions in the CAI modules assigned to Group 1.

Virtual Slides

The use of VS or static images of pathology showed no significant difference in their effect on immediate or subsequent student exam performance. However, it should be noted that students were only tested with static images, whether or not they were assigned to a CAI learning module that incorporated VS. Thus, it is possible that results would be different if students had been tested with VS. Several other studies have reported student acceptance of learning Pathology from virtual slides instead of glass slides and microscopes (Anyanwu, Agu, & Anyaehie, 2012; Kim et al., 2008; Krippendorf & Lough, 2005; Kumar et al., 2004; Nivala et al., 2012; Pantanowitz, 2012).

However, only two reported studies were found in the literature that compared student exam performance after learning with glass slides or VS (Anyanwu et al., 2012; Kumar et al., 2004). Kumar, et al (2004), showed that performance with VS was identical or minimally improved compared to historical controls. Anyanwu (Anyanwu et al., 2012) found that students learning from virtual slides performed significantly better in an examination using VS and students showed a significantly higher preference for VS. However, no reported studies have rigorously compared the learning of pathology with VS vs. static images.

Theories of adult learning based on constructivist principles would suggest that the use of virtual slides would add context to the case-based nature of the learning exercise. Images are fragmented snapshots of tissue pathology that do not allow one to see how those fragments relate to one another to produce a composite picture of the pathology and its occurrence in adjacent more normal tissue. A student learning from a virtual slide is more likely to recognize a high-power image of the pathology than a student who learns from a set of high-power images and is then tested with a virtual slide. Virtual slides are a true simulation of the way that pathologic diagnoses are made in the course of patient care, adding realism and relevance to the case-based learning exercise.

On the other hand, review of a virtual slide takes more time and cognitive investment than review of a static image, even if annotated. In one study in which web-based pathology cases and images were used in place of glass slides and microscopes, students showed 10% greater attendance at lab exercises using the web-based, self-study tutorials and reported more satisfaction in learning from them, but did not tend to use the zoom and scroll options that were available to enhance learning from the digital images, suggesting that students made an effort to avoid the added cognitive complexity involved with zooming and scrolling of the images (Marchevsky, Relan, & Baillie, 2003). Mayer and Moreno (Mayer & Moreno, 2003) point out that there is a potential negative effect in complex multimedia presentations that can divide a learner's attention with overloaded multimedia formats. Students who are poor visual learners may have more difficulty with a virtual slide than a straight-forward static image of pathology. Grunwald cautions that CAI learning modules must accommodate these abilities and limitations in learners, and so underscores the importance of testing CAI modules that may seem intuitively beneficial, but prove to be less effective in practice (Grunwald & Corsbie-Massay, 2006).

Interactive Image Annotations

The use of annotations on the virtual slides and images in this study was appreciated by students, as reflected in their survey comments. The annotations were used in an effort to decrease cognitive load. Annotations can direct attention by placing text in contiguity with the image in an effort to decrease the load on working memory (Sweller et al., 1998; Mayer, 2001; Issa et al., 2011). Annotations are more efficient and less time-consuming for learning, and so, provide the student with more time to process the information into working memory (Tabbers et al., 2004). Tabbers, et al (2004) found that adding visual cues to pictures in a class room-administered CAI learning module resulted in higher scores on a post-test of learning for Year-2 students in education. However, one possible detractor may be that less time is spent in seeing the pathology illustrated by the annotation in the context of other features apparent in the image, which could inhibit the ability of the student to recognize the features in future images of the pathology.

Interactivity in the use of the image annotations is supported by theories of learning that predict student engagement and control of learning will improve student satisfaction and exam performance, building on theories of iterative learning and problem solving, as well as making students more active and more in control of their learning (Argyris & Schon, 1974; Issa et al., 2011; Kolb, 1984; Mayer, 2005). However, evaluation of the effect of interactive annotations on exam performance by multiple regression analysis showed that interactive annotations had a weak negative correlation with student scores on PostTest 1. It may be that the annotation interactivity distracted students from the text-image association or that the use of it was confounded by the need to use it within the time constraints of the in-class lab exercise.

Guiding Questions

The use of guiding questions in the CAI learning modules was cited as particularly helpful by students in open-ended survey comments, but failed to show efficacy in multiple regression analysis of student Post Test exam scores. The use of quiding questions is well supported by learning theories that suggest that guiding questions draw attention to the material to be learned and promote student reflection and self-explanation, which will help students to integrate new information with existing information to create a new understanding that can deepen learning and strengthen long-term memory (Bude et al., 2012; Chi et al., 2013; Kirschner et al., 2006; Mayer, 2004). Like the interactive annotations, the failure to show a positive effect of guiding questions on exam scores in this lab exercise may be related to the design of the lab exercise. All CAI learning modules presented all of the required information, so it does appear that students who learned from CAI modules without guiding questions were able to absorb and retain that information as well as students who were provided with guiding questions. It may be that the time constraints of the lab exercise did not allow for the sort of reflection and self-explanation that was intended.

Student Class Rank

It was not anticipated that class rank would have a significant impact on PostTest performance of students working with the CAI learning modules, but the failure of the learning interventions to affect student posttest performance raised a concern that the learning interventions might have differential effects on students dependent on their general knowledge base and aptitude for learning pathology, as reflected in the student's level of high or low class rank at the time of the CAI lab exercise. Kalyuga (2010) had reported that hyperlinks could have a differential effect on students with high- or low-level knowledge. However, the results of this study showed no such differential effect of the learning interventions depending on class rank, but found class rank itself to be a predictor of exam performance on questions directed to knowledge gained after working with the CAI learning module. When analyzed as a continuous variable with multiple linear regression, student class rank was a strong predictor of outcome for PostTest 1, and both high and low quintile rank were weak predictors of performance on PostTests 2 and 3. This outcome is perhaps not surprising, since all learning begins with an interaction between the information to be learned, which is impacted by the difficulty of the topic and the volume of information, and the capability, previous experience and knowledge of the learner to process it. These are features of intrinsic cognitive load (Plass et al., 2010).

The results in this case suggest that the difficulty of the topic, volume of information in time allotted, or format of presentation was too great in this case to allow students with weaker skills in learning pathology from text and images to overcome their deficiencies. It is possible that having more time and a more relaxed atmosphere for review of the materials would have elicited a different result.

Survey Results

Survey results of student satisfaction and perceived learning in use of the CAI modules did show, as predicted, positive correlations with the case-based nature of the exercise and with some features of the learning interventions, including: the hyperlinks that allowed all of the necessary learning information about the cases to be included within the CAI module; the slide label annotations, but not necessarily their interactivity; and the use of guiding questions. It is perhaps noteworthy that positive and negative themes that came out of student comments were not elicited by direct questions concerning any of the learning interventions, since not all students were exposed to each intervention. All student comments about features of the CAI learning modules were anonymous and spontaneous. The feature of case-based learning that the students appreciated is supported by theory, which suggests that students are more motivated to learn and lessons have greater impact when presented in a real-world context to enhance relevance of the lessons for the student (Eva, MacDonald, Rodenburg, & Regehr, 2000; Cook et al., 2010b).

The use of hyperlinks that students identified as a positive aspect of the CAI modules is supported by the higher level of interactivity they stimulate, which can be associated with deeper learning (Ross & Tuovinen, 2001). The use of hyperlinks also fosters cognitive flexibility and the appreciation that higher level learning is not linear, but occurs in a matrix of information (Graddy, 2001; Spiro et al., 1992). In this study, the hyperlinks were clearly identified as to whether they were sources of basic or more advanced information, since it is known that the use of hyperlinks to additional advanced information on a topic may increase extraneous cognitive load for a student with low-level knowledge of the topic; whereas, hyperlinks to basic information may help a student with low-level knowledge but inhibit learning in a student with greater knowledge of the topic, referred to as "expertise reversal effect" (Kalyuga, 2010).

The other features that students found positive are reflected in a study of their use in a Pathology course and in several reviews of multimedia and CAI learning modules that integrate multimedia learning theories with reports on student satisfaction and performance after the use of these programs (Reid et al., 2000). Guidelines have been established on the basis of these reports that underscore the student-identified positive features of the CAI learning module used in this study (Eva et al., 2000; Grunwald & Corsbie-Massay, 2006).

Survey results of student satisfaction and perceived learning in use of the CAI modules also showed, as predicted, negative correlations with technical difficulties a student might have experienced in the smooth operation of the program on their personal computers or stress concerning the time-limited, in-class setting or the inability to interact in discussion with other students about the cases.

Study Limitations

In considering the delimitations of the experimental design used in the current study, it is important to consider what limitations the study design might have on measurable outcomes. The lack of positive effect of the learning interventions on exam performance in the current study is surprising in consideration of relevant learning theories and reports in the literature, which suggest that these interventions should decrease cognitive load and improve student learning in use of the CAI learning modules (Cook et al., 2010; Clark & Feldon, 2005; Eva et al., 2000; Grunwald & Corsbie-Massay, 2006). It is quite possible, however, that limitations imposed by the experimental design and implementation may have affected the outcome. Tabbers, et al (2004) also found unexpected results most likely related to experimental design. In their study, visual text cues to images were replaced with spoken text within an in-class CAI learning module. Dual-channel multimedia theory would suggest that the change would reduce cognitive load and improve learning and retention when auditory and visual channels were employed and synchronized (Mayer, 2001). However, the authors found that posttest scores were actually lower when auditory cues were added to the program. The authors attributed the unexpected outcome to the experimental design, which allowed students to pace their own learning, in contrast to previous experiments that had used system-paced instruction in the experimental design.

Several limitations in the design of the current quantitative randomized-controlled experimental study of a CAI learning module for pathology education may have impacted on the study outcomes and might have some impact on future studies of this nature. One such limitation was the setting for student use of the CAI learning module in an in-class, time-delimited, computer-based laboratory exercise in which an entire class of students worked independently, Although the setting was logistically necessary in order to provide the controlled environment necessary for the current experimental research design, it may not have been optimal for all students. Some personal and laboratory computers were not as efficient as others, which may have increased frustration for some students in use of the program and decreased the ease of access to information provided in hyperlinks.

Time constraints may have been an additional disincentive in use of the hyperlinks for some students. Time constraints may also have affected learning from virtual slides, which often take more time to review than a static image. The use of interactive features and guiding questions also take more time for self-reflection and integration of information. Although time for the laboratory exercise should have been adequate, as determined by usability testing and previous experience with use of the case studies involved in other Pathology course exercises, some students may have been stressed by the time delimited nature of the exercise, especially with the spector of a posttest exam for credit that was to be administered at the end of the laboratory exercise.

The posttest exams may themselves have been a source of limitation in interpretation of results. In the current research study, the PostTests 1-3 were not designed to measure an improvement in the basic knowledge that was measured by the PreTest. The PostTests 1-3 were designed to measure the ability of students to apply basic knowledge to the interpretation of pathology presented in clinical cases, similar to the applications they learned about in the CAI learning module; the CAI lab exercise was their only in-class opportunity to learn that information before being tested on it at the end of the laboratory exercise. Some students may have felt undue pressure and anxiety during the laboratory exercise, knowing that a test would follow; especially if their preparation using the pre-session materials had been weak. Student performance on

PostTests 2 and 3 may also have been affected by anxiety and the distraction of learning new information in other subjects that was being tested concurrently.

The inability to collaborate that was a requirement for the controlled nature of this study also proved to be a frustration for some students. Collaboration can be very helpful to learning for some students. Eva, et al (2000) pointed out that the inability for students to collaborate can weaken learning from CAI modules and that the design of experiments to test such modules may show no differences between CAI modules in randomized-controlled studies when this important variable is eliminated.

One unexpected finding in the current study was the strong effect of class rank on tests of exam performance associated with the CAI. Class rank was not considered in the original study design, but the variable was added to determine if the learning interventions might have a differential effect dependent on a student's basic knowledge and aptitude for learning pathology, as determined by class rank. There were no significant differences that could be attributed to the learning interventions when they were studied within quintile levels of class rank; however, class rank itself was a strong predictor of student performance on posttest exams.

Implications for Future Research and Practice

The findings of this study have implications for future research and practice in the study of learning interventions in CAI learning modules for studies of pathology. In future research designs, the learning task should be limited to one or two variables for study in use of the CAI learning module. If pre-session materials are required, the amount of study should be limited and focused to just what is required to learn the new information presented in the CAI learning module. In this way, students should not be overwhelmed by the learning task presented in class and should require less time to complete the exercise, which should in turn decrease stress and maximize learning. Testing anxiety should be minimized as much as possible, but exam performance is likely to be a very important outcome for measurement. One possible approach may be to make the exercise and/ or the posttest optional, but with opportunities for extra credit.

A reduced number of variables for study would also allow the experimental design to include fewer students in the study, while maintaining adequate power in the experimental design to test the significance of the results. Fewer students in the research design might also make it possible to consider venues for testing that have optimal technological support and that are practical and conducive to student learning.
A small number of students for testing may allow the exercise to take place in different settings and at times that could be more individualized for optimal student learning.

The significant impact of student class rank on exam performance must be taken into account in future research studies that test learning interventions in the instructional design of CAI learning modules and that use exam performance as a measure of student learning. In considering the design of quantitative, randomized and controlled experimental studies of CAI learning modules for pathology education, it is important that these studies minimize the effect of a student's aptitude, as estimated by class rank, on exam performance as a measure of learning after use of a CAI learning intervention. This might be done by including sufficiently large numbers of students at each quintile of class rank, such that the power is sufficient to determine statistical significance of the CAI learning intervention. Alternatively, cross-over studies might be a more practical approach, in which an entire class experiences the CAI learning module with and without the intervention, such that all students are exposed to all variables in the study in a way that can eliminate any advantage or disadvantage the learning intervention have on exam results that might effect a student's grade in the course.

Finally, the student survey in the current study added information about the use of the CAI learning modules that is important to understand. Students were eager to give an opinion about what was helpful or distracting in the use of the CAI. Student feedback can shape the development of a CAI program for future use and point to features that should be further developed or rejected. Student satisfaction with use of the CAI learning modules also suggests features that would make it more likely for a student to self-select such a program for use in the future.

Significance of Research Findings and Recommendations

Medical education is currently undergoing major changes in the U.S. with the implementation of new curricula that that markedly decrease lecture and class time in the preclinical years and increase student self-study time. Students are provided with learning resources online and expected to come to classes that are designed to apply the information that they have learned (Cooke et al., 2010). There is an ever-increasing number of Computer Assisted Instruction learning modules available to medical students, often with an instructional design based on adult learning principles; however, there are very few studies that test the effectiveness of the instructional design on measurable outcomes (Cook, 2009).

This study was undertaken as a direct response to the call for research on principles of multimedia design in medical education (Association of American Medical Colleges, 2007); and in particular, the need to focus research questions on the effectiveness of interventions in design rather than comparisons of CAI with other modalities of education (Cook, 2009; Berman et al., 2008). Cook (Cook et al., 2008) found large positive effects of internet-based instruction compared to no intervention, but variable and small effects compared to traditional methods. Other reviews of CAI effectiveness have arrived at similar findings (Greenhalgh, 2001; Chumley-Jones et al., 2002).

This study is the first randomized-controlled, quantitative study to directly compare the three learning interventions tested - use of virtual slides, interactive annotations and guiding questions, in web-based, case-based, self-study, CAI learning modules employed in medical student learning of pathology. The results indicate that these learning interventions are likely to be perceived as helpful, in light of student comments on a satisfaction survey in the current study, but are not likely to have a positive impact on student performance or retention when provided to second-year medical students work independently with the CAI during an in-class laboratory exercise in a Pathology course, as determined by posttest measurements of applied knowledge, similar to those in the current study.

The variable that did have a strong effect on posttesting was class rank, which most likely reflects a student's background of knowledge and aptitude in pathology, the subject area of the CAI. It seems likely that students of higher education, such as medical students, are able in most cases to adapt their learning strategy to obtain the information they need from a CAI learning module to meet the objectives required given the context in which the CAI is presented. For these students, the most important essential features of instructional design in learning with the use of a CAI learning program are likely to be the quality of the informational content, targeting to an appropriate level of learning, adjustment of the volume of information to well-focused objectives and delivery of the CAI in a computer- or web-based environment that is user friendly to a spectrum of different computer systems and browsers, so that all students find easy access to the available features of the program.

If the essential features are met in a CAI learning program, the learning interventions may then help or hinder the learning process by their effects on the efficiency of learning what is needed at the student's level of knowledge and experience. Interactive or static annotations may be a help to the learning of a student inexperienced with images of pathology, but they may be a time-consuming distraction to a student with more advanced learner. Virtual slides can be an ideal way to learn pathology when there is time for a student to interact with the slides and guidance is provided in interpretation of the pathology, but virtual slides can be a frustrating and inefficient way for inexperienced students to learn the essential features of pathology if there is not adequate guidance and little time available for their study. The positive impact of learning interventions to enhance medical student learning in CAI may be as much about increasing efficiency, ease of use and satisfaction with the learning process, as it is about increasing the amount of information learned.

It seems likely that CAI learning modules will play an increasing role in the future training and continuing education of physicians. The quality of the content in a CAI learning modules and the delivery of content that is appropriate to the stage of the learner will, of course, be expected to impact learning. However, it may be that specific learning interventions similar to the ones tested here may not have as much of an impact at the higher level of education in medical school as other factors that make learning easy or convenient. It may be difficult to demonstrate the impact of a learning intervention on exam performance.

As important as quantitative research is in finding evidence for the effectiveness of CAI learning programs and the use of learning interventions in their instructional design, however, the results of the satisfaction survey in the current study and similar studies indicate that student satisfaction and perception of learning can be important and perhaps critical factors in student learning, even though they may be difficult to quantify. The positive themes that came out of open-ended student comments in the current study point to several attributes of the CAI learning modules that were perceived as helpful to student learning, including the ability to control the pace of learning, interactivity with the program, use of annotations, guiding questions and self-selection of informational content through hyperlinks in a way that makes all of the required information on a topic accessible in one place. These features can be important incentives for students to use such programs. Student-directed and life-long learning requires students to self-select their own learning opportunities and make use of them. If the CAI learning module can provide that opportunity, students are likely to choose it on their own. There is no hope of learning, if students choose not to access the information to be learned.

REFERENCE LIST

The Cambridge Handbook of Multimedia Learning. (2005). New York, NY: Cambridge University Press.

Aamodt, A. & Plaza, E. (1994). Case-based reasoning. *Artifical Intelligence Communnications, 7,* 39-59.

Anyanwu, G. E., Agu, A. U., & Anyaehie, U. B. (2012). Enhancing learning objectives by use of simple virtual microscopic slides in cellular physiology and histology: impact and attitudes. *Advances in Physiology Education, 36,* 158-163.

Argyris, C. & Schon, D. A. (1974). *Theory in practice: Increasing professional effectiveness*. San Francisco,CA: Jossey-Bass.

Argyris, C. & Schon, D. A. (1978). *Organizational learning: A theory of action perspective.* Reading, MA: Addison Wesley.

Argyris, C. & Schon, D. A. (1996). *Organizational learning II: Theory, method and practice.* Reading, MA: Addison Wesley.

Association of American Medical Colleges (2007). *Use of Educational Technology in Medical Education.* Washington, DC.

Barrows, H. S. (1994). *Practice-Based Learning. Problem-based learning applied to medical education.* Springfield, IL: Southern Illinois University School of Medicine.

Berman, N. B., Fall, L. H., Maloney, C. G., & Levine, D. A. (2008). Computer-assisted instruction in clinical education: a roadmap to increasing CAI implementation. *Advances in Health Sciences Education, 13,* 373-383.

Blake, H. (2010). Computer-based learning objects in healthcare: the student experience. *International Journal of Nursing Education Scholarship, 7,* Article.

Bloom, B. S. (1994). *Reflections on the development and use of the taxonomy.* Chicago, IL: University of Chocago Press.

Bruner, J. (1960). *The process of education.* Cambridge, MA: Harcard University Press.

Bude, L., van de Wiel, M. W., Imbos, T., & Berger, M. P. (2012). The effect of guiding questions on students' performance and attitude towards statistics. *British Journal of Educational Psychology, 82,* 340-359.

Case, S. M. & Swanson, D. B. (1998). *Constructing written test questions for the basic and clinical sciences.* (3rd edition ed.) Philadelphia, PA: National Board of Medical Examiners.

Castillo-Paige, L. (2012). Diversity in Medical Education: Facts & Figures 2012 [Computer software]. Washington, D.C.: Diversity Policy and Programs, Association of American Medical Colleges.

Centre for Excellence in Teaching and Learning in Resuable Learning Objects (RLO-CETL) (2005). Learning Object Questionnaire - Use of the Learning Object. Centre for Excellence in Teaching and Learning in Resuable Learning Objects

(RLO-CETL) Tool-kit, Creative Commons NonCommercial-ShareAlike 2.5 License (CC BY-NC-SA 2.5) [On-line]. Available: http://www.rlo-cetl.ac.uk

Chi, M. T. H., Bassock, M., Lewis, M., Reimann, P., & Glaser, R. (2013). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13,* 145-182.

Chumley-Jones, H. S., Dobbie, A., & Alford, C. L. (2002). Web-based learning: sound educational method or hype? A review of the evaluation literature. [Review] [39 refs]. *Academic Medicine, 77,* S86-S93.

Clark, R. E. & Feldon, D. F. (2005). Five Common but Questionable Principles of Multimedia Learning. In R.E.Mayer (Ed.),

Cambridge Handbook of Multimedia Learning. (pp. 97-116). New York, N.Y.: Cambridge University Press.

Cook, D. A. (2009). The failure of e-learning research to inform educational practice, and what we can do about it. *Medical Teacher, 31,* 158-162.

Cook, D. A., Levinson, A. J., Garside, S., Dupras, D. M., ., Erwin, P. J. et al. (2008). Internet-based learning in the health professions: a meta-analysis. *Journal of the American Medical Association, 300,* 1181-1196.

Cook, D. A., Levinson, A. J., & Garside, S. (2010a). Time and learning efficiency in Internet-based learning: a systematic review and meta-analysis. *Advances in Health Sciences Education, 15,* 755-770.

Cook, D. A., Levinson, A. J., Garside, S., Dupras, D. M., Erwin, P. J., & Montori, V. M. (2010b). Instructional design variations in internet-based learning for health professions education: a systematic review and meta-analysis. *Academic Medicine, 85,* 909-922.

Cooke, M., Irby, D. M., & O'Brien, C. (2010). *Educating physicians: A call for reform of medical school and residency.* San Francisco, CA: Jossey-Bass.

Creswell, J. W. (2003). *Research design. Qualitative, quantitative and mixed methods approaches, 2nd ed.* Thousand Oaks, CA: SAGE Publications, Inc.

Crotty, M. (2005). *The foundations of social research. Meaning and perspective in the research process.* Thousand Oaks, CA: SAGE Publications.

Eva, K. W., MacDonald, R. D., Rodenburg, D., & Regehr, G. (2000). Maintaining the Characteristics of Effective Clinical Teachers in Computer Assisted Learning Environments. *Advances in Health Sciences Education, 5,* 233-246.

Gagne, R., Briggs, L., & Wager, W. (1992). *Principles of instructional design. 4th edition.* Fort Worth, TX: HBJ College Publishers.

Gardenfors, P. & Johansson, P. (2005). *Cognition, Education and Communication Technology.* Mahwah, N.J.: Lawrence Erlbaum Associates.

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology, 95,* 393-408.

Graddy, D. B. (2001). Cognitive flexibility theory as a pedagogy for web-based course design. Indiana - Purdue University Fort Wayne [On-line]. Available: http://www.ipfw.edu/as/tohe/2001/Papers/graddy/graddy.htm

Greenhalgh, T. (2001). Computer assisted learning in undergraduate medical education. *British Medical Journal, 322,* 40-44.

Grubbs, F. (1969). Procedures for Detecting Outlying Observations in Samples. Technometrics 11[1], 1-21.

Ref Type: Journal (Full)

Grunwald, T. & Corsbie-Massay, C. (2006). Guidelines for cognitively efficient multimedia learning tools: educational strategies, cognitive load, and interface design. *Academic Medicine, 81,* 213-223.

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). *Applied Statistics for the Behavioral Sciences.* (fifth ed.) New York, N.Y.: Houghton Mifflin Co.

Issa, N., Schuller, M., Santacaterina, S., Shapiro, M., Wang, E., Mayer, R. E. et al. (2011). Applying multimedia design principles enhances learning in medical education. *Medical Education, 45,* 818-826.

Jha, V. & Duffy, S. (2002). 'Ten golden rules' for designing software in medical education: results from a formative evaluation of DIALOG. *Medical Teacher, 24,* 417-421.

Kalyuga, S. (2010). Cognitive Load Theory. In J.L.Plass, R. Moreno, & R. Brunken (Eds.), (pp. 48-64). New York, NY: Cambridge University Press.

Kim, M. H., Park, Y., Seo, D., Lim, Y. J., Kim, D., Kim, C. W. et al. (2008). Virtual microscopy as a practical alternative to conventional microscopy in pathology education. Basic and Applied Pathology 1, 46-48.

Ref Type: Journal (Full)

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist,

Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. Educational Psychologist 41, 75-86.

Ref Type: Journal (Full)

Knowles, M. S., Holton, E. F., & Swanson, R. A. (1998). *The adult learner. The definitive classic in adult education and human resource development. (5th edition)*. Houston, TX: Gulf Publishing Co.

Kolb, D. A. (1984). *Experiential Learning: Experience as the source of learning and development.* Englewood Cliffs, N.J.: Prentice Hall.

Krippendorf, B. B. & Lough, J. (2005). Complete and rapid switch from light microscopy to virtual microscopy for teaching medical histology. *Anatomical Record (Part B: The New Anatomist), 285,* 19-25.

Kumar, R. K., Velan, G. M., Korell, S. O., Kandara, M., Dee, F. R., & Wakefield, D. (2004). Virtual microscopy for learning and assessment in pathology. *Journal of Pathology, 204,* 613-618.

Lave, J. & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation.* New York, NY: Cambridge University Press.

Malachowski, M. (2002). ADDIE based five-step method towards instructional design. [On-line]. Available: http://fog.ccsf.cc.ca.us/~mmalacho/OnLine/ADDIE.html

Marchevsky, A. M., Relan, A., & Baillie, S. (2003). Self-instructional "virtual pathology" laboratories using web-based technology enhance medical school teaching of pathology. *Human Pathology, 34,* 423-429.

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist, 32,* 1-19.

Mayer, R. E. (2001). *The Cambridge Handbook of Multimedia Learning.* New York, NY: Cambridge University Press.

Mayer, R. E. (2004). Should There Be a Three-Strikes Rule Against Pure Discovery Learning? The Case for Guided Methods of Instruction. *American Psychologist, 59,* 14-19.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R.E.Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 19-30). New York, N.Y.: Cambridge University Press.

Mayer, R. E. & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38,* 43-52.

Nivala, M., Saljo, R., Rystedt, H., Kronqvist, P., & Lehtinen, E. (2012). Using virtual microscopy to scaffold learning of pathology: a naturalistic experiment on the role of visual and conceptual cues. Instructional Science 40, 799-811.

Ref Type: Journal (Full)

Pantanowitz, L. (2012). Whole Slide Imaging for Educational Purposes. *Journal of Pathology Informatics, 3,* 46.

Plass, J. L., Moreno, R., & Brunken, R. (2010). *Cognitive Load Theory*. (First edition ed.) New York, N.Y.: Cambridge University Press.

Reid, W. A., Harvey, J., Watson, G. R., Luqmani, R., Harkin, P. J., & Arends, M. J. (2000). Medical student appraisal of interactive computer-assisted learning programs embedded in a general pathology course. *Journal of Pathology, 191,* 462-465.

Ross, G. C. & Tuovinen, J. E. (2001). Deep versus surface learning with multimedia in nursing education development and evaluation of WoundCare. *Computers in Nursing, 19,* 213-223.

Savery, J. R. & Duffy, T. M. (1995). Problem based learning: An instructional model and its constructivist framework. *Educational Technology, 35,* 31-38.

Schon, D. A. (1983). *The Reflective Practitioner. How professionals think in action.* London: Temple Smith.

Spiro, R. J., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L. (1992). Cognitive flexibility, constructivism and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured doamins. In T.Duffy & D. Jonassen (Eds.), *Constructivism and the Technology of Instruction.* (Hillsdale, NJ: Erlbaum.

Spiro, R. J. & Jehng, J. (1990). Cognitive flexibility and hypertext: Theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In D.Nix & R. Spiro (Eds.), *Cognition, Education and Multimedia.* (Hillsdale, NJ: Erlbaum.

Sweller, J. (1988). Cognitive Load During Problem-Solving - Effects on Learning. *Cognitive Science, 12,* 257-285.

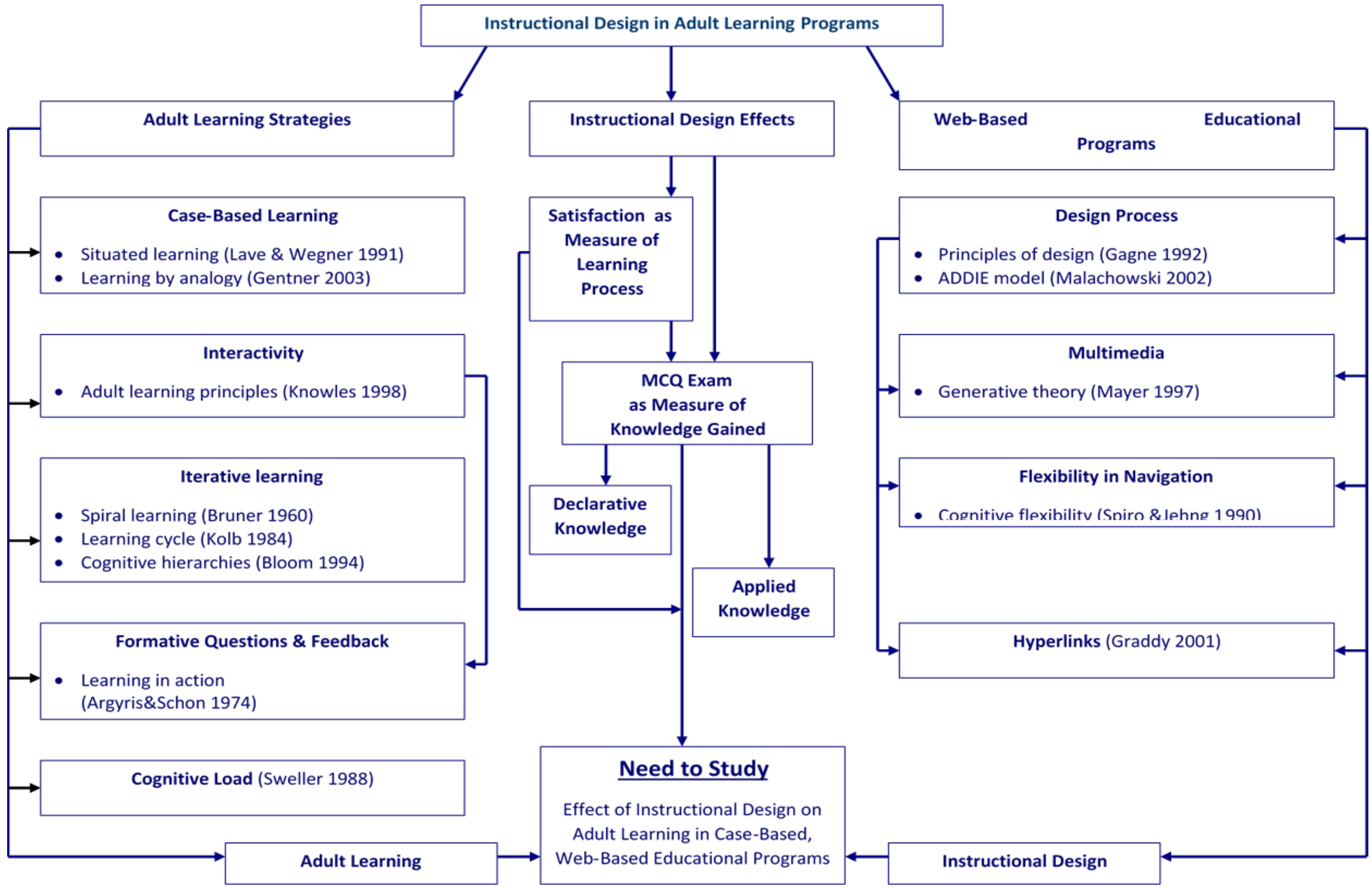
Sweller, J., Merrienboer, J., & Paas, F. (1998). Cognitive architecture and Instructional design. *Educational Psychology Review, 10,* 251-296.

Tabbers, H. K., Martens, R. L., & Van Merrienboer, J. J. (2004). Multimedia instructions and cognitive load theory: effects of modality and cueing. *British Journal of Educational Psychology, 74,* 71-81.

Thalheimer, W. (2003). The learning benefits of questions. Retrieved November 23, 2012, from http://www.work-learning.com/ma/PP_WP003.asp [On-line]. Available: http://www.work-learning.com/ma/PP_WP003.asp

Van Merrienboer, J. J. & Sweller, J. (2010). Cognitive load theory in health professional education: design principles and strategies. *Medical Education, 44,* 85-93.

APPENDIX A – LITERATURE REVIEW MAP



APPENDIX B

SURVEY INSTRUMENT

Student Survey 02/11/13 **Computer-Assisted Instruction (CAI) Liver Case Studies** **Survey is anonymous**

	Please take a few minutes to provide your feedback on the Computer-Assisted Instruction (CAI) modules on liver disease.						
	What was your group number (required)						
	Name (not required - but helpful)						
	Score	**1**	**2**	**3**	**4**	**5**	**6**
	Directions: **Please indicate your level of agreement with the statements.**	**Agree Strongly**	**Agree Moderately**	**Agree Slightly**	**Disagree Slightly**	**Disagree Moderately**	**Disagree Strongly**
1	Clear about purpose and objectives						
2	Easy to navigate						
3	Introduced new information clearly						
4	Would recommend to my peers						
5	Would like to use similar CAI to learn pathology for other diseases						
6	Learning was set at the right level for my stage in the course						
7	The CAI complemented other teaching in the Pathology course						
8	I enjoyed learning from the CAI						
9	I learned better from the CAI than from lab and small group						
10	I learned better from the CAI than from virtual slides and PP cases when I study on my own						
11	What was very good/ helpful about the CAI?						
12	What was very bad/ distracting about the CAI?						
13	What suggestions do you have to improve the CAI?						

APPENDIX C

Appendix C

Numbers of Students Responding to Open-Ended Questions in the Survey Questionnaire

Groups	Total Respondents N	Question 1: What was very good/helpful about the CAI? N	Question 2: What was very bad/distracting about the CAI? N	Question 3: What suggestions do you have to improve the CAI? N	Additional Comments N
1	20	18	18	14	5
2	21	15	15	14	6
3	19	16	15	11	6
4	17	14	15	12	6
5	21	19	19	10	8
6	14	9	11	9	3
7	17	11	11	10	4
8	15	13	13	10	5
Total Responses	144	115	117	90	43

APPENDIX D

STATISTICAL TESTING CONDITIONS AND ASSUMPTIONS

Between Student Group Scores on PreTest Analyzed by One-Way ANOVA

Results for each of the eight student groups on the PreTest were analyzed by one-way ANOVA.

There were no extreme outliers, as determined by Grubb's test at an alpha level < .01 (Grubns, 1969).. Data was normally distributed, as determined by skewness and kurtosis at an alpha level < .01 ($z < \pm 2.58$), with exception of Group 2 with a skewness of -0.961 (*SE* = 0.491) and kurtosis of -0.732 (*SE* = 0.033) and Group 7 with a skewness of -0.238 (*SE* = 0.491) and kurtosis of -0.773 (*SE* = 0.035). Homogeneity of variance was equal for student groups, as assessed by Levene's Test of Homogeneity of Variance ($p = .602$). The alpha level for the ANOVA result was set at $p < .05$.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
PreT	Between Groups	.108	7	.015	.467	.858
	Within Groups	5.455	165	.033		
	Total	5.563	172			

Between Student Group Scores on PostTest 1-3 MCQ Directed to Case A (without learning interventions) Analyzed by One-Way ANOVA

Results were analyzed by one-way ANOVA. There were no extreme outliers. Data was normally distributed as determined by skewness and kurtosis at an alpha level < .01 (z < ±2.58), with exception of three groups for PostTest 1 and one group for PostTest 2, as follows: PostTest 1 Group 1 with a skewness of -2.531 (*SE* = 0.481) and kurtosis of 3.855 (*SE* = 0.935), Group 4 with a skewness of -2.157 (*SE* = 0.501) and kurtosis of -3.539 (*SE* = 0.972) and Group 5 with a skewness of -1.596 (*SE* = 0.501) and kurtosis of 1.896 (*SE* = 0.972). The assumption of homogeneity was also violated, as assessed by Levene's Test of Homogeneity of Variance (HOV) in PostTest 1 (*p* = <.001), but not in PostTest 2 (*p* = .996) or PostTest 3 (*p* = .784). Thus, Welch's ANOVA test was used for determinations of significance at an alpha level of < .05.

ANOVA

PostTest1short_3Q_Case A	Between Groups	.447	7	.064	1.485	.176
	Within Groups	7.093	165	.043		
	Total	7.540	172			
PostTest2short_3Q_Case A	Between Groups	.330	7	.047	.919	.493
	Within Groups	8.461	165	.051		
	Total	8.791	172			
PostTest3short_4Q_Case A	Between Groups	.257	7	.037	.503	.831
	Within Groups	12.047	165	.073		
	Total	12.304	172			

Between Group Student Exam Scores for Questions directed to Cases B and C (with learning interventions) in PostTests 1-3 Analyzed by One-Way ANOVA

Results were analyzed by one-way ANOVA. There were no extreme outliers, as assessed by Grubb's test at an alpha level < .01 , with exception of one outlier in Group 4 PostTest 1 on MCQ directed to Cases B-C. This student score was not deleted, since the outcome was unchanged when the score was eliminated. Data was normally distributed as determined by skewness and kurtosis at an alpha level < .01 ($z < \pm 2.58$), with exception of Group 4 in PostTest 1 with a skewness of -1.473 (*SE* = 0.501) and kurtosis of 3.974 (*SE* = 0.972) and Group 6 with a skewness of 0.444 (*SE* = 0.512) and kurtosis of 0.648 (*SE* = 0.043). Homogeneity of variances was equal for all CAI-modules in 8 student groups, as assessed by Levene's Test of Homogeneity of Variance, PostTest 1 (p= .731), PostTest 2 (p= .656) and PostTest 3 (p= .594).

ANOVA

PostTest2difference_Case BC	Between Groups	.312	7	.045	1.185	.314
	Within Groups	6.198	165	.038		
	Total	6.509	172			
PostTest2difference_Case BC	Between Groups	.312	7	.045	1.185	.314
	Within Groups	6.198	165	.038		
	Total	6.509	172			
PostTest3difference_Case BC	Between Groups	.184	7	.026	.836	.559
	Within Groups	5.176	165	.031		
	Total	5.360	172			

Between Group Student Exam Scores for PreTest and PostTest 1-3 Exam Results with PostTest Results directed to Case A (no learning interventions) or to Cases B-C (with learning interventions) Analyzed by Kruskall-Wallis Non-parametric Testing

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of PreT is the same across categories of Grp.	Independent-Samples Kruskal-Wallis Test	.882	Retain the null hypothesis.
2	The distribution of PostTest1short_3Q_Case A is the same across categories of Grp.	Independent-Samples Kruskal-Wallis Test	.146	Retain the null hypothesis.
3	The distribution of PostTest1difference_Case BC is the same across categories of Grp.	Independent-Samples Kruskal-Wallis Test	.245	Retain the null hypothesis.
4	The distribution of PostTest2short_3Q_Case A is the same across categories of Grp.	Independent-Samples Kruskal-Wallis Test	.612	Retain the null hypothesis.
5	The distribution of PostTest2difference_Case BC is the same across categories of Grp.	Independent-Samples Kruskal-Wallis Test	.245	Retain the null hypothesis.
6	The distribution of PostTest3short_4Q_Case A is the same across categories of Grp.	Independent-Samples Kruskal-Wallis Test	.818	Retain the null hypothesis.
7	The distribution of PostTest3difference_Case BC is the same across categories of Grp.	Independent-Samples Kruskal-Wallis Test	.467	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Comparison of Student Exam Performance on MCQ Directed to Case A (without learning interventions) or Cases B-C (with learning interventions) using Paired Samples t-test

There were no extreme outliers and the group scores were normally distributed, as determined by skewness and kurtosis at an alpha level < .01 (z < ±2.58), with exception of those described above in the results of ANOVA testing.

Group 1

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	PostTest1short_3Q_Case A - PostTest1difference_Case BC	.19563	.19275	.04019	.11228	.27898	4.868	22	.000
Pair 2	PostTest2short_3Q_Case A - PostTest2difference_Case BC	-.04349	.37615	.07843	-.20615	.11917	-.555	22	.585
Pair 3	PostTest3short_4Q_Case A - PostTest3difference_Case BC	-.01631	.34398	.07173	-.16506	.13244	-.227	22	.822

Group 2

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	PostTest1short_3Q_Case A - PostTest1difference_Case BC	.16853	.23412	.04991	.06473	.27233	3.376	21	.003
Pair 2	PostTest2short_3Q_Case A - PostTest2difference_Case BC	.12499	.26665	.05685	.00677	.24322	2.199	21	.039
Pair 3	PostTest3short_4Q_Case A - PostTest3difference_Case BC	.03977	.25702	.05480	-.07418	.15373	.726	21	.476

Group 3

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	PostTest1short_3Q_Case A - PostTest1difference_Case BC	.12497	.28521	.06081	-.00148	.25143	2.055	21	.053
Pair 2	PostTest2short_3Q_Case A - PostTest2difference_Case BC	.11174	.29166	.06218	-.01757	.24106	1.797	21	.087
Pair 3	PostTest3short_4Q_Case A - PostTest3difference_Case BC	.01158	.30587	.06521	-.12404	.14719	.178	21	.861

Group 4

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	PostTest1short_3Q_Case A - PostTest1difference_Case BC	.22617	.15565	.03397	.15532	.29702	6.659	20	.000
Pair 2	PostTest2short_3Q_Case A - PostTest2difference_Case BC	.15079	.24524	.05352	.03915	.26242	2.818	20	.011
Pair 3	PostTest3short_4Q_Case A - PostTest3difference_Case BC	.06547	.24564	.05360	-.04634	.17728	1.221	20	.236

Group 5

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	PostTest1short_3Q_Case A - PostTest1difference_Case BC	.14482	.26634	.05812	.02358	.26606	2.492	20	.022
Pair 2	PostTest2short_3Q_Case A - PostTest2difference_Case BC	.10118	.33086	.07220	-.04943	.25179	1.401	20	.176
Pair 3	PostTest3short_4Q_Case A - PostTest3difference_Case BC	-.11905	.24836	.05420	-.23210	-.00600	-2.197	20	.040

Group 6

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	PostTest1short_3Q_Case A - PostTest1difference_Case BC	.07707	.25444	.05689	-.04201	.19615	1.355	19	.191
Pair 2	PostTest2short_3Q_Case A - PostTest2difference_Case BC	.01666	.29812	.06666	-.12287	.15618	.250	19	.805
Pair 3	PostTest3short_4Q_Case A - PostTest3difference_Case BC	.00000	.29801	.06664	-.13947	.13947	.000	19	1.000

Group 7

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	PostTest1short_3Q_Case A - PostTest1difference_Case BC	.20076	.18706	.03988	.11782	.28370	5.034	21	.000
Pair 2	PostTest2short_3Q_Case A - PostTest2difference_Case BC	.10605	.25675	.05474	-.00779	.21989	1.937	21	.066
Pair 3	PostTest3short_4Q_Case A - PostTest3difference_Case BC	-.11932	.32385	.06905	-.26291	.02427	-1.728	21	.099

Group 8

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	PostTest1short_3Q_Case A - PostTest1difference_Case BC	.20452	.24394	.05201	.09636	.31267	3.932	21	.001
Pair 2	PostTest2short_3Q_Case A - PostTest2difference_Case BC	.14772	.29651	.06322	.01625	.27918	2.337	21	.029
Pair 3	PostTest3short_4Q_Case A - PostTest3difference_Case BC	-.00568	.27407	.05843	-.12719	.11583	-.097	21	.923

Within Student Group Sequential Scores on PostTest 1-3 MCQ Directed at Case A Analyzed by one-way repeated measures ANOVA with pairwise comparisons

There were no extreme outliers, but there were some test results that did not show a normal distribution for some groups in PostTests 1 and 2, as determined by skewness and kurtosis at an alpha level < .01 ($z < \pm 2.58$) as noted above. The assumption of sphericity was also violated, as assessed by Mauchly's test of sphericity, $\chi^2(2) = 6.10$, $p = .047$. Therefore, a Greenhouse-Geisser correction was applied ($\varepsilon = 0.966$).

Group 1-Case A

Estimates

Measure: PT

Time	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	.765	.036	.690	.841
2	.957	.024	.907	1.006
3	.696	.055	.581	.810
4	.511	.058	.391	.631

Pairwise Comparisons

Measure: PT

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.^{b}	95% Confidence Interval for Difference^b	
					Lower Bound	Upper Bound
1	2	-.191*	.037	.000	-.299	-.084
	3	.070	.057	1.000	-.095	.234
	4	.254*	.067	.006	.059	.449
2	1	.191*	.037	.000	.084	.299
	3	.261*	.063	.002	.079	.443
	4	.446*	.059	.000	.273	.618
3	1	-.070	.057	1.000	-.234	.095
	2	-.261*	.063	.002	-.443	-.079
	4	.185	.076	.139	-.035	.404
4	1	-.254*	.067	.006	-.449	-.059
	2	-.446*	.059	.000	-.618	-.273
	3	-.185	.076	.139	-.404	.035

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Group 2-Case A

Estimates

Measure: PT

Time	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	.718	.049	.617	.819
2	.833	.048	.734	.933
3	.727	.052	.619	.836
4	.500	.055	.387	.613

Pairwise Comparisons

Measure: PT

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.^{b}	95% Confidence Interval for Difference^b	
					Lower Bound	Upper Bound
1	2	-.115	.053	.242	-.269	.038
	3	-.009	.077	1.000	-.233	.215
	4	.218*	.074	.045	.003	.433
2	1	.115	.053	.242	-.038	.269
	3	.106	.067	.776	-.090	.302
	4	.333*	.082	.003	.095	.572
3	1	.009	.077	1.000	-.215	.233
	2	-.106	.067	.776	-.302	.090
	4	.227*	.059	.006	.054	.400
4	1	-.218*	.074	.045	-.433	-.003
	2	-.333*	.082	.003	-.572	-.095
	3	-.227*	.059	.006	-.400	-.054

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Group 3-Case A

Estimates

Measure: PT

Time	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	.765	.036	.690	.841
2	.957	.024	.907	1.006
3	.696	.055	.581	.810
4	.511	.058	.391	.631

Pairwise Comparisons

Measure: PT

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b] Lower Bound	95% Confidence Interval for Difference[b] Upper Bound
1	2	-.191*	.037	.000	-.299	-.084
	3	.070	.057	1.000	-.095	.234
	4	.254*	.067	.006	.059	.449
2	1	.191*	.037	.000	.084	.299
	3	.261*	.063	.002	.079	.443
	4	.446*	.059	.000	.273	.618
3	1	-.070	.057	1.000	-.234	.095
	2	-.261*	.063	.002	-.443	-.079
	4	.185	.076	.139	-.035	.404
4	1	-.254*	.067	.006	-.449	-.059
	2	-.446*	.059	.000	-.618	-.273
	3	-.185	.076	.139	-.404	.035

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Group 4-Case A

Estimates

Measure: PT

Time	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	.786	.037	.709	.862
2	.905	.047	.807	1.002
3	.794	.049	.692	.895
4	.524	.062	.395	.653

Pairwise Comparisons

Measure: PT

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b] Lower Bound	95% Confidence Interval for Difference[b] Upper Bound
1	2	-.119*	.044	.013	-.211	-.027
	3	-.008	.061	.897	-.135	.119
	4	.262*	.066	.001	.125	.399
2	1	.119*	.044	.013	.027	.211
	3	.111*	.053	.049	.000	.222
	4	.381*	.062	.000	.252	.510
3	1	.008	.061	.897	-.119	.135
	2	-.111*	.053	.049	-.222	.000
	4	.270*	.064	.000	.136	.404
4	1	-.262*	.066	.001	-.399	-.125
	2	-.381*	.062	.000	-.510	-.252
	3	-.270*	.064	.000	-.404	-.136

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Group 5-Case A

Estimates

Measure: PT

Time	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	.790	.039	.709	.872
2	.889	.042	.801	.976
3	.810	.043	.719	.900
4	.440	.049	.339	.542

Pairwise Comparisons

Measure: PT

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b] Lower Bound	95% Confidence Interval for Difference[b] Upper Bound
1	2	-.098	.051	.406	-.248	.051
	3	-.019	.067	1.000	-.216	.178
	4	.350*	.042	.000	.226	.474
2	1	.098	.051	.406	-.051	.248
	3	.079	.051	.810	-.070	.228
	4	.448*	.066	.000	.255	.642
3	1	.019	.067	1.000	-.178	.216
	2	-.079	.051	.810	-.228	.070
	4	.369*	.062	.000	.189	.549
4	1	-.350*	.042	.000	-.474	-.226
	2	-.448*	.066	.000	-.642	-.255
	3	-.369*	.062	.000	-.549	-.189

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Case 6-Case A

Estimates

Measure: PT

Time	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	.755	.039	.674	.836
2	.783	.061	.657	.910
3	.717	.056	.600	.833
4	.488	.059	.365	.610

Pairwise Comparisons

Measure: PT

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b] Lower Bound	95% Confidence Interval for Difference[b] Upper Bound
1	2	-.028	.057	.623	-.147	.090
	3	.038	.066	.568	-.100	.176
	4	.267*	.062	.000	.138	.397
2	1	.028	.057	.623	-.090	.147
	3	.067	.082	.428	-.106	.239
	4	.296*	.083	.002	.121	.470
3	1	-.038	.066	.568	-.176	.100
	2	-.067	.082	.428	-.239	.106
	4	.229*	.083	.013	.055	.403
4	1	-.267*	.062	.000	-.397	-.138
	2	-.296*	.083	.002	-.470	-.121
	3	-.229*	.083	.013	-.403	-.055

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Group 7-Case A

Estimates

Measure: PT

Time	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	.736	.042	.648	.825
2	.848	.036	.773	.924
3	.788	.041	.702	.874
4	.398	.063	.267	.529

Pairwise Comparisons

Measure: PT

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.b	95% Confidence Interval for Difference[b] Lower Bound	Upper Bound
1	2	-.112*	.048	.029	-.212	-.012
	3	-.052	.040	.214	-.135	.032
	4	.339*	.056	.000	.222	.455
2	1	.112*	.048	.029	.012	.212
	3	.061	.042	.162	-.026	.148
	4	.451*	.056	.000	.334	.567
3	1	.052	.040	.214	-.032	.135
	2	-.061	.042	.162	-.148	.026
	4	.390*	.058	.000	.269	.512
4	1	-.339*	.056	.000	-.455	-.222
	2	-.451*	.056	.000	-.567	-.334
	3	-.390*	.058	.000	-.512	-.269

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Group 8-Case A

Estimates

Measure: PT

Time	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	.777	.035	.704	.851
2	.864	.042	.776	.951
3	.818	.042	.730	.906
4	.489	.065	.354	.623

Pairwise Comparisons

Measure: PT

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.b	95% Confidence Interval for Difference[b] Lower Bound	Upper Bound
1	2	-.086	.052	.113	-.195	.022
	3	-.041	.053	.453	-.152	.070
	4	.289*	.057	.000	.169	.408
2	1	.086	.052	.113	-.022	.195
	3	.045	.059	.451	-.078	.169
	4	.375*	.078	.000	.213	.537
3	1	.041	.053	.453	-.070	.152
	2	-.045	.059	.451	-.169	.078
	4	.330*	.081	.001	.161	.498
4	1	-.289*	.057	.000	-.408	-.169
	2	-.375*	.078	.000	-.537	-.213
	3	-.330*	.081	.001	-.498	-.161

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Within group serial student exam performance on MCQ Directed to Case A in PostTests 1-3 Analyzed by one-way repeated measures ANOVA with pairwise comparisons

There were no extreme outliers, except the one student identified above in Group 4 for a test result in PostTest 1 MCQ directed to Cases B-C. There was a normal distribution of data for all PostTests 1-3, as determined by skewness and kurtosis at an alpha level < .01 ($z < \pm 2.58$). The assumption of sphericity was not violated, as assessed by Mauchly's test of sphericity, $\chi^2(2) = .244$, $p = .885$.

Group 1-Case B-C

Estimates

Measure: PT

Time	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	.765	.036	.690	.841
2	.761	.030	.698	.824
3	.739	.041	.655	.824
4	.527	.034	.456	.598

Pairwise Comparisons

Measure: PT

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b] Lower Bound	95% Confidence Interval for Difference[b] Upper Bound
1	2	.004	.036	1.000	-.101	.110
	3	.026	.055	1.000	-.132	.184
	4	.238*	.040	.000	.121	.355
2	1	-.004	.036	1.000	-.110	.101
	3	.022	.044	1.000	-.104	.148
	4	.234*	.047	.000	.098	.369
3	1	-.026	.055	1.000	-.184	.132
	2	-.022	.044	1.000	-.148	.104
	4	.212*	.054	.004	.057	.367
4	1	-.238*	.040	.000	-.355	-.121
	2	-.234*	.047	.000	-.369	-.098
	3	-.212*	.054	.004	-.367	-.057

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Group 2-Case B-C

Estimates

Measure: PT

Time	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	.718	.049	.617	.819
2	.665	.029	.605	.725
3	.602	.035	.530	.674
4	.460	.027	.405	.515

Pairwise Comparisons

Measure: PT

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b] Lower Bound	95% Confidence Interval for Difference[b] Upper Bound
1	2	.053	.043	1.000	-.071	.177
	3	.116	.063	.488	-.068	.300
	4	.258*	.060	.002	.083	.433
2	1	-.053	.043	1.000	-.177	.071
	3	.063	.048	1.000	-.077	.202
	4	.205*	.038	.000	.093	.316
3	1	-.116	.063	.488	-.300	.068
	2	-.063	.048	1.000	-.202	.077
	4	.142*	.043	.020	.017	.267
4	1	-.258*	.060	.002	-.433	-.083
	2	-.205*	.038	.000	-.316	-.093
	3	-.142*	.043	.020	-.267	-.017

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Group 3-Case B-C

Estimates

Measure: PT

Time	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	.732	.033	.663	.801
2	.693	.029	.632	.754
3	.631	.045	.538	.724
4	.477	.034	.407	.547

Pairwise Comparisons

Measure: PT

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b] Lower Bound	Upper Bound
1	2	.039	.042	1.000	-.083	.160
	3	.101	.046	.241	-.033	.236
	4	.255*	.042	.000	.134	.376
2	1	-.039	.042	1.000	-.160	.083
	3	.063	.046	1.000	-.070	.195
	4	.216*	.044	.000	.087	.346
3	1	-.101	.046	.241	-.236	.033
	2	-.063	.046	1.000	-.195	.070
	4	.154*	.050	.034	.008	.299
4	1	-.255*	.042	.000	-.376	-.134
	2	-.216*	.044	.000	-.346	-.087
	3	-.154*	.050	.034	-.299	-.008

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Group 4-Case B-C

Estimates

Measure: PT

Time	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	.786	.037	.709	.862
2	.679	.038	.599	.758
3	.643	.041	.558	.728
4	.458	.041	.373	.543

Pairwise Comparisons

Measure: PT

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b] Lower Bound	Upper Bound
1	2	.107	.037	.052	-.001	.215
	3	.143*	.048	.046	.002	.284
	4	.327*	.055	.000	.165	.489
2	1	-.107	.037	.052	-.215	.001
	3	.036	.047	1.000	-.103	.174
	4	.220*	.056	.005	.057	.384
3	1	-.143*	.048	.046	-.284	-.002
	2	-.036	.047	1.000	-.174	.103
	4	.185*	.036	.000	.079	.291
4	1	-.327*	.055	.000	-.489	-.165
	2	-.220*	.056	.005	-.384	-.057
	3	-.185*	.036	.000	-.291	-.079

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Group 5-Case B-C

Estimates

Measure: PT

Time	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	.790	.039	.709	.872
2	.744	.034	.673	.815
3	.708	.048	.608	.809
4	.560	.043	.470	.649

Pairwise Comparisons

Measure: PT

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b] Lower Bound	Upper Bound
1	2	.046	.050	1.000	-.100	.193
	3	.082	.048	.607	-.058	.222
	4	.231*	.045	.000	.099	.363
2	1	-.046	.050	1.000	-.193	.100
	3	.036	.061	1.000	-.143	.214
	4	.185*	.051	.011	.034	.335
3	1	-.082	.048	.607	-.222	.058
	2	-.036	.061	1.000	-.214	.143
	4	.149*	.051	.049	.000	.297
4	1	-.231*	.045	.000	-.363	-.099
	2	-.185*	.051	.011	-.335	-.034
	3	-.149*	.051	.049	-.297	.000

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Group 6-Case B-C

Estimates

Measure: PT

Time	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	.755	.039	.674	.836
2	.706	.028	.648	.764
3	.700	.034	.628	.772
4	.488	.040	.403	.572

Pairwise Comparisons

Measure: PT

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b] Lower Bound	Upper Bound
1	2	.049	.034	1.000	-.051	.149
	3	.055	.042	1.000	-.070	.180
	4	.267*	.052	.000	.115	.420
2	1	-.049	.034	1.000	-.149	.051
	3	.006	.039	1.000	-.108	.121
	4	.219*	.038	.000	.106	.332
3	1	-.055	.042	1.000	-.180	.070
	2	-.006	.039	1.000	-.121	.108
	4	.213*	.045	.001	.079	.346
4	1	-.267*	.052	.000	-.420	-.115
	2	-.219*	.038	.000	-.332	-.106
	3	-.213*	.045	.001	-.346	-.079

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Group 7-Case B-C

Estimates

Measure: PT

Time	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	.736	.042	.648	.825
2	.648	.043	.558	.738
3	.682	.039	.600	.763
4	.517	.041	.433	.601

Pairwise Comparisons

Measure: PT

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
1	2	.089	.049	.526	-.055	.233
	3	.055	.042	1.000	-.067	.176
	4	.219*	.049	.001	.078	.361
2	1	-.089	.049	.526	-.233	.055
	3	-.034	.049	1.000	-.176	.108
	4	.131*	.044	.043	.003	.259
3	1	-.055	.042	1.000	-.176	.067
	2	.034	.049	1.000	-.108	.176
	4	.165*	.046	.011	.031	.299
4	1	-.219*	.049	.001	-.361	-.078
	2	-.131*	.044	.043	-.259	-.003
	3	-.165*	.046	.011	-.299	-.031

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Group 8-Case B-C

Estimates

Measure: PT

Time	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	.777	.035	.704	.851
2	.659	.032	.592	.726
3	.670	.048	.571	.770
4	.494	.044	.403	.586

Pairwise Comparisons

Measure: PT

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
1	2	.118*	.037	.028	.009	.227
	3	.107	.043	.130	-.019	.232
	4	.283*	.047	.000	.148	.418
2	1	-.118*	.037	.028	-.227	-.009
	3	-.011	.041	1.000	-.131	.108
	4	.165*	.047	.012	.029	.301
3	1	-.107	.043	.130	-.232	.019
	2	.011	.041	1.000	-.108	.131
	4	.176*	.053	.020	.022	.331
4	1	-.283*	.047	.000	-.418	-.148
	2	-.165*	.047	.012	-.301	-.029
	3	-.176*	.053	.020	-.331	-.022

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Comparison of Performance Scores Between High-, Mid- and Low=Quintiles of Class Rank on PreTest and PostTest 1-3 Analyzed by one-way ANOVA

Results analyzed by one-way ANOVA showed two outliers, as determined by Grubb's test at an alpha level < .01, except for two student scores for PostTest 1 with MCQ directed to Case A, one student score in the mid-quintile group and one in the high-quintile group. The two student scores were left in the analysis. Data was normally distributed as determined by skewness and kurtosis at an alpha level < .01 ($z < \pm 2.58$), with exception of scores on PostTest 1 with MCQ directed to Case A for High and Middle Quintile groups, with skewness -2.768 (*SE* = .398) and kurtosis 7.646 (*SE* = .778) for the high-quintile group and with skewness -1.437 (*SE* = .236) and kurtosis 1.637 (*SE* = .467) for the mid-quintile group. The mid-quintile group also showed deviation from normal on PostTest 2 with MCQ directed to Case A with skewness -.735 (*SE* = .236) and kurtosis .251 (*SE* = .467) . The scores were left in the analysis unchanged, since ANOVA is robust to minor deviations in normality (Hinkle et al., 2003). Homogeneity of variances was equal for Quintile groups in all PostTest results, as assessed by Levene's Test of Homogeneity of Variance, except for MCQ directed to Case A (without learning interventions) in PostTest 1 ($p < .001$). Thus, Welch's ANOVA was used for analysis of statistical significance.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
PreT	Between Groups	1.018	2	.509	19.026	.000
	Within Groups	4.546	170	.027		
	Total	5.563	172			
PostTest1short_3Q_Case A	Between Groups	.331	2	.166	3.906	.022
	Within Groups	7.209	170	.042		
	Total	7.540	172			
PostTest1long_11Q__Case ABC	Between Groups	.449	2	.224	13.301	.000
	Within Groups	2.868	170	.017		
	Total	3.317	172			
PostTest1difference_Case BC	Between Groups	.513	2	.257	11.739	.000
	Within Groups	3.717	170	.022		
	Total	4.230	172			
PostTest2short_3Q_Case A	Between Groups	.290	2	.145	2.898	.058
	Within Groups	8.501	170	.050		
	Total	8.791	172			
PostTest2long_11Q_Case ABC	Between Groups	1.422	2	.711	45.578	.000
	Within Groups	2.652	170	.016		
	Total	4.074	172			
PostTest2difference_Case BC	Between Groups	2.175	2	1.087	42.644	.000
	Within Groups	4.335	170	.025		
	Total	6.509	172			
PostTest3short_4Q_Case A	Between Groups	1.650	2	.825	13.163	.000
	Within Groups	10.654	170	.063		
	Total	12.304	172			
PostTest3long_12Q_Case ABC	Between Groups	.996	2	.498	24.580	.000
	Within Groups	3.445	170	.020		
	Total	4.441	172			
PostTest3difference_Case BC	Between Groups	.768	2	.384	14.219	.000
	Within Groups	4.592	170	.027		
	Total	5.360	172			

Significant Differences and Confidence Intervals are as follows:

PostTest 1-Case A between high and low quintile scores 13.4% (95% CI 2.9 to 24.0), *p*= .009
PostTest 1-Case A between high and middle quintile scores 8.9% (95% CI 0.9 to 16.8), *p*= .025

PostTest 3-Cases A between high and low quintile scores 26.9% (95% CI 12.9 to 41.1), *p*= <.001
PostTest 3-Cases A between high and middle quintile scores 23.1% (95% CI 11.9 to 34.2), *p*= <.001

PostTest 1-Cases B-C between high and low quintile scores 17.3% (95% CI 8.87 to 25.72), *p*= <.001
PostTest 1-Cases B-C between high and middle quintile scores 7.3% (95% CI 1.3 to 13.2), *p*= .013
PostTest 1-Cases B-C between low and middle quintile scores 10.0% (95% CI 0.9 to 16.8), *p*= .025

PostTest 2-Cases B-C between high and low quintile scores 35.8% (95% CI 27.2 to 44.3), *p*= <.001
PostTest 2-Cases B-C between high and middle quintile scores 17.5% (95% CI 10.4 to 24.6), *p*= <.001
PostTest 2-Cases B-C between low and middle quintile scores 18.2% (95% CI 10.9 to 25.6), *p*= <.001

PostTest 3-Cases B-C between high and low quintile scores 20.9% (95% CI 11.5 to 30.4), *p*= <.001
PostTest 3-Cases B-C between high and middle quintile scores 12.6% (95% CI 4.4 to 20.8), *p*= . 001
PostTest 3-Cases B-C between low and middle quintile scores 8.3% (95% CI 1.0 to 15.6), *p*= . 021

Correlation of Class Rank as a Continuous Variable with Exam Performance Scores on PostTests 1 and 2 Analyzed by Pearson's correlation coefficient.

Correlations

		PostTest1difference_Case BC	CumAve	PreT	PostTest1short_3Q_Case A	PostTest2short_3Q_Case A	PostTest2difference_Case BC	PostTest3short_4Q_Case A	PostTest3difference_Case BC
PostTest1difference_Case BC	Pearson Correlation	1	.290**	.353**	.232**	.019	.250**	.224**	.193*
	Sig. (2-tailed)		.000	.000	.002	.802	.001	.003	.011
	N	173	173	173	173	173	173	173	173
CumAve	Pearson Correlation	.290**	1	.467**	.269**	.181*	.644**	.388**	.407**
	Sig. (2-tailed)	.000		.000	.000	.017	.000	.000	.000
	N	173	173	173	173	173	173	173	173
PreT	Pearson Correlation	.353**	.467**	1	.348**	.062	.268**	.254**	.203**
	Sig. (2-tailed)	.000	.000		.000	.419	.000	.001	.007
	N	173	173	173	173	173	173	173	173
PostTest1short_3Q_Case A	Pearson Correlation	.232**	.269**	.348**	1	.138	.159*	.045	.077
	Sig. (2-tailed)	.002	.000	.000		.070	.037	.553	.314
	N	173	173	173	173	173	173	173	173
PostTest2short_3Q_Case A	Pearson Correlation	.019	.181*	.062	.138	1	-.008	.104	.102
	Sig. (2-tailed)	.802	.017	.419	.070		.919	.174	.180
	N	173	173	173	173	173	173	173	173
PostTest2difference_Case BC	Pearson Correlation	.250**	.644**	.268**	.159*	-.008	1	.252**	.305**
	Sig. (2-tailed)	.001	.000	.000	.037	.919		.001	.000
	N	173	173	173	173	173	173	173	173
PostTest3short_4Q_Case A	Pearson Correlation	.224**	.388**	.254**	.045	.104	.252**	1	.192*
	Sig. (2-tailed)	.003	.000	.001	.553	.174	.001		.012
	N	173	173	173	173	173	173	173	173
PostTest3difference_Case BC	Pearson Correlation	.193*	.407**	.203**	.077	.102	.305**	.192*	1
	Sig. (2-tailed)	.011	.000	.007	.314	.180	.000	.012	
	N	173	173	173	173	173	173	173	173

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Predictive Ability of Learning Interventions and Class Rank (High and Low Quintile) for Questions directed to Cases B-C (with learning interventions) in PostTest 1-3 Analyzed by Multiple Regression

PostTest 1-Cases B-C

	Mean	Std. Deviation	N
PostTest1 difference_Case BC	.6944	.15682	173
HiQuintile	.20	.403	173
LoQuintile	.19	.394	173
MidQuintile	.61	.490	173
VirtualSlides	.49	.501	173
InterAnnotate	.50	.501	173
LeadQues	.50	.513	173

Correlations

		PostTest1 difference_Case BC	HiQuintile	LoQuintile	MidQuintile	VirtualSlides	InterAnnotate	LeadQues
Pearson Correlation	PostTest1 difference_Case BC	1.000	.248	-.298	.035	-.039	-.160	-.115
	HiQuintile	.248	1.000	-.245	-.626	.023	.040	-.067
	LoQuintile	-.298	-.245	1.000	-.603	-.036	.012	-.012
	MidQuintile	.035	-.626	-.603	1.000	.010	-.043	.065
	VirtualSlides	-.039	.023	-.036	.010	1.000	.029	-.006
	InterAnnotate	-.160	.040	.012	-.043	.029	1.000	-.006
	LeadQues	-.115	-.067	-.012	.065	-.006	-.006	1.000
Sig. (1-tailed)	PostTest1 difference_Case BC	.	.000	.000	.322	.307	.018	.066
	HiQuintile	.000	.	.001	.000	.381	.299	.189
	LoQuintile	.000	.001	.	.000	.320	.438	.440
	MidQuintile	.322	.000	.000	.	.450	.289	.198
	VirtualSlides	.307	.381	.320	.450	.	.352	.470
	InterAnnotate	.018	.299	.438	.289	.352	.	.471
	LeadQues	.066	.189	.440	.198	.470	.471	.
N	PostTest1 difference_Case BC	173	173	173	173	173	173	173
	HiQuintile	173	173	173	173	173	173	173
	LoQuintile	173	173	173	173	173	173	173
	MidQuintile	173	173	173	173	173	173	173
	VirtualSlides	173	173	173	173	173	173	173
	InterAnnotate	173	173	173	173	173	173	173
	LeadQues	173	173	173	173	173	173	173

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.685	5	.137	6.450	.000[b]
	Residual	3.545	167	.021		
	Total	4.230	172			

a. Dependent Variable: PostTest1 difference_Case BC

b. Predictors: (Constant), LeadQues, InterAnnotate, LoQuintile, VirtualSlides, HiQuintile

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	.748	.024		31.517	.000	.701	.795					
	HiQuintile	.073	.029	.187	2.548	.012	.016	.129	.248	.193	.180	.933	1.071
	LoQuintile	-.101	.029	-.253	-3.459	.001	-.158	-.043	-.298	-.259	-.245	.938	1.066
	VirtualSlides	-.015	.022	-.048	-.673	.502	-.059	.029	-.039	-.052	-.048	.998	1.002
	InterAnnotate	-.051	.022	-.163	-2.303	.023	-.095	-.007	-.160	-.175	-.163	.997	1.003
	LeadQues	-.033	.022	-.106	-1.499	.136	-.075	.010	-.115	-.115	-.106	.995	1.005

a. Dependent Variable: PostTest1 difference_Case BC

Note. LeadQues are Guiding Questions

PostTest 2-Cases B-C

Descriptive Statistics

	Mean	Std. Deviation	N
PostTest2difference_Case BC	.6720	.19454	173
HiQuintile	.20	.403	173
LoQuintile	.19	.394	173
VirtualSlides	.49	.501	173
InterAnnotate	.50	.501	173
LeadQues	.50	.513	173

Correlations

		PostTest2difference_Case BC	HiQuintile	LoQuintile	VirtualSlides	InterAnnotate	LeadQues
Pearson Correlation	PostTest2difference_Case BC	1.000	.453	-.459	.090	-.080	-.104
	HiQuintile	.453	1.000	-.245	.023	.040	-.067
	LoQuintile	-.459	-.245	1.000	-.036	.012	-.012
	VirtualSlides	.090	.023	-.036	1.000	.029	-.006
	InterAnnotate	-.080	.040	.012	.029	1.000	-.006
	LeadQues	-.104	-.067	-.012	-.006	-.006	1.000
Sig. (1-tailed)	PostTest2difference_Case BC	.	.000	.000	.120	.149	.086
	HiQuintile	.000	.	.001	.381	.299	.189
	LoQuintile	.000	.001	.	.320	.438	.440
	VirtualSlides	.120	.381	.320	.	.352	.470
	InterAnnotate	.149	.299	.438	.352	.	.471
	LeadQues	.086	.189	.440	.470	.471	.
N	PostTest2difference_Case BC	173	173	173	173	173	173
	HiQuintile	173	173	173	173	173	173
	LoQuintile	173	173	173	173	173	173
	VirtualSlides	173	173	173	173	173	173
	InterAnnotate	173	173	173	173	173	173
	LeadQues	173	173	173	173	173	173

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2.306	5	.461	18.326	.000[b]
	Residual	4.203	167	.025		
	Total	6.509	172			

a. Dependent Variable: PostTest2difference_Case BC

b. Predictors: (Constant), LeadQues, InterAnnotate, LoQuintile, VirtualSlides, HiQuintile

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	.692	.026		26.776	.000					
	HiQuintile	.173	.031	.359	5.581	.000	.453	.396	.347	.933	1.071
	LoQuintile	-.182	.032	-.368	-5.739	.000	-.459	-.406	-.357	.938	1.066
	VirtualSlides	.027	.024	.071	1.134	.258	.090	.087	.071	.998	1.002
	InterAnnotate	-.036	.024	-.092	-1.481	.140	-.080	-.114	-.092	.997	1.003
	LeadQues	-.032	.024	-.084	-1.354	.178	-.104	-.104	-.084	.995	1.005

a. Dependent Variable: PostTest2difference_Case BC

Note. LeadQues are Guiding Questions

PostTest 3-Cases B-C

Descriptive Statistics

	Mean	Std. Deviation	N
PostTest3difference_Case BC	.4978	.17652	173
HiQuintile	.20	.403	173
LoQuintile	.19	.394	173
VirtualSlides	.49	.501	173
InterAnnotate	.50	.501	173
LeadQues	.50	.513	173

Correlations

		PostTest3difference_Case BC	HiQuintile	LoQuintile	VirtualSlides	InterAnnotate	LeadQues
Pearson Correlation	PostTest3difference_Case BC	1.000	.333	-.256	.094	-.062	-.140
	HiQuintile	.333	1.000	-.245	.023	.040	-.067
	LoQuintile	-.256	-.245	1.000	-.036	.012	-.012
	VirtualSlides	.094	.023	-.036	1.000	.029	-.006
	InterAnnotate	-.062	.040	.012	.029	1.000	-.006
	LeadQues	-.140	-.067	-.012	-.006	-.006	1.000
Sig. (1-tailed)	PostTest3difference_Case BC	.	.000	.000	.108	.210	.033
	HiQuintile	.000	.	.001	.381	.299	.189
	LoQuintile	.000	.001	.	.320	.438	.440
	VirtualSlides	.108	.381	.320	.	.352	.470
	InterAnnotate	.210	.299	.438	.352	.	.471
	LeadQues	.033	.189	.440	.470	.471	.
N	PostTest3difference_Case BC	173	173	173	173	173	173
	HiQuintile	173	173	173	173	173	173
	LoQuintile	173	173	173	173	173	173
	VirtualSlides	173	173	173	173	173	173
	InterAnnotate	173	173	173	173	173	173
	LeadQues	173	173	173	173	173	173

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.914	5	.183	6.864	.000[b]
	Residual	4.446	167	.027		
	Total	5.360	172			

a. Dependent Variable: PostTest3difference_Case BC

b. Predictors: (Constant), LeadQues, InterAnnotate, LoQuintile, VirtualSlides, HiQuintile

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	.509	.027		19.133	.000					
	HiQuintile	.123	.032	.281	3.850	.000	.333	.286	.271	.933	1.071
	LoQuintile	-.083	.033	-.184	-2.535	.012	-.256	-.193	-.179	.938	1.066
	VirtualSlides	.029	.025	.083	1.172	.243	.094	.090	.083	.998	1.002
	InterAnnotate	-.026	.025	-.074	-1.047	.297	-.062	-.081	-.074	.997	1.003
	LeadQues	-.043	.024	-.124	-1.748	.082	-.140	-.134	-.123	.995	1.005

a. Dependent Variable: PostTest3difference_Case BC

Note. LeadQues are Guiding Questions

Predictive Ability of Learning Interventions and Class Rank (Cum Ave) as a Continuous Variable for Student Scores on MCQ directed to Cases -C in PostTests 1-3 Analyzed by Multiple Regression

There was one student score that was an outlier, as described above for one-way ANOVA between groups, but this score was left in the analysis. The assumptions of linearity, independence of errors, homoscedasticity, unusual points and normality of residuals were met in all case.

PostTest 1-Cases B-C

Descriptive Statistics

	Mean	Std. Deviation	N
PostTest1 difference_Case BC	.6944	.15682	173
CumAve	.8073	.07601	173
VirtualSlides	.49	.501	173
InterAnnotate	.50	.501	173
LeadQues	.50	.513	173

Correlations

		PostTest1 difference_Case BC	CumAve	VirtualSlides	InterAnnotate	LeadQues
Pearson Correlation	PostTest1 difference_Case BC	1.000	.290	-.039	-.160	-.115
	CumAve	.290	1.000	.042	.026	-.014
	VirtualSlides	-.039	.042	1.000	.029	-.006
	InterAnnotate	-.160	.026	.029	1.000	-.006
	LeadQues	-.115	-.014	-.006	-.006	1.000
Sig. (1-tailed)	PostTest1 difference_Case BC	.	.000	.307	.018	.066
	CumAve	.000	.	.291	.366	.425
	VirtualSlides	.307	.291	.	.352	.470
	InterAnnotate	.018	.366	.352	.	.471
	LeadQues	.066	.425	.470	.471	.
N	PostTest1 difference_Case BC	173	173	173	173	173
	CumAve	173	173	173	173	173
	VirtualSlides	173	173	173	173	173
	InterAnnotate	173	173	173	173	173
	LeadQues	173	173	173	173	173

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.536	4	.134	6.088	.000[b]
	Residual	3.694	168	.022		
	Total	4.230	172			

a. Dependent Variable: PostTest1 difference_Case BC

b. Predictors: (Constant), LeadQues, InterAnnotate, VirtualSlides, CumAve

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	.254	.122		2.090	.038	.014	.495					
	CumAve	.608	.149	.294	4.079	.000	.314	.902	.290	.300	.294	.997	1.003
	VirtualSlides	-.015	.023	-.047	-.647	.518	-.059	.030	-.039	-.050	-.047	.997	1.003
	InterAnnotate	-.052	.023	-.167	-2.310	.022	-.097	-.008	-.160	-.175	-.167	.999	1.001
	LeadQues	-.034	.022	-.112	-1.551	.123	-.078	.009	-.115	-.119	-.112	1.000	1.000

a. Dependent Variable: PostTest1 difference_Case BC

Note. LeadQues refers to Guiding Questions.

PostTest 2-Cases B-C

Descriptive Statistics

	Mean	Std. Deviation	N
PostTest2difference_Case BC	.6720	.19454	173
CumAve	.8073	.07601	173
VirtualSlides	.49	.501	173
InterAnnotate	.50	.501	173
LeadQues	.50	.513	173

Correlations

		PostTest2difference_Case BC	CumAve	VirtualSlides	InterAnnotate	LeadQues
Pearson Correlation	PostTest2difference_Case BC	1.000	-.030	.090	-.080	-.104
	CumAve	-.030	1.000	-.060	-.014	.148
	VirtualSlides	.090	-.060	1.000	.029	-.006
	InterAnnotate	-.080	-.014	.029	1.000	-.006
	LeadQues	-.104	.148	-.006	-.006	1.000
Sig. (1-tailed)	PostTest2difference_Case BC	.	.349	.120	.149	.086
	CumAve	.349	.	.215	.425	.026
	VirtualSlides	.120	.215	.	.352	.470
	InterAnnotate	.149	.425	.352	.	.471
	LeadQues	.086	.026	.470	.471	.
N	PostTest2difference_Case BC	173	173	173	173	173
	CumAve	173	173	173	173	173
	VirtualSlides	173	173	173	173	173
	InterAnnotate	173	173	173	173	173
	LeadQues	173	173	173	173	173

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.168	4	.042	1.113	.352[b]
	Residual	6.341	168	.038		
	Total	6.509	172			

a. Dependent Variable: PostTest2difference_Case BC

b. Predictors: (Constant), LeadQues, InterAnnotate, VirtualSlides, CumAve

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	.711	.161		4.419	.000	.393	1.029					
	CumAve	-.026	.197	-.010	-.132	.895	-.416	.364	-.030	-.010	-.010	.974	1.026
	VirtualSlides	.035	.030	.091	1.194	.234	-.023	.094	.090	.092	.091	.996	1.004
	InterAnnotate	-.032	.030	-.083	-1.089	.278	-.091	.026	-.080	-.084	-.083	.999	1.001
	LeadQues	-.039	.029	-.103	-1.334	.184	-.097	.019	-.104	-.102	-.102	.978	1.022

a. Dependent Variable: PostTest2difference_Case BC

Note. LeadQues are Guiding Questions

PostTest 3-Cases B-C

Descriptive Statistics

	Mean	Std. Deviation	N
PostTest3difference_Case BC	.4978	.17652	173
CumAve	.8073	.07601	173
VirtualSlides	.49	.501	173
InterAnnotate	.50	.501	173
LeadQues	.50	.513	173

Correlations

		PostTest3difference_Case BC	CumAve	VirtualSlides	InterAnnotate	LeadQues
Pearson Correlation	PostTest3difference_Case BC	1.000	.036	.094	-.062	-.140
	CumAve	.036	1.000	-.060	-.014	.148
	VirtualSlides	.094	-.060	1.000	.029	-.006
	InterAnnotate	-.062	-.014	.029	1.000	-.006
	LeadQues	-.140	.148	-.006	-.006	1.000
Sig. (1-tailed)	PostTest3difference_Case BC	.	.320	.108	.210	.033
	CumAve	.320	.	.215	.425	.026
	VirtualSlides	.108	.215	.	.352	.470
	InterAnnotate	.210	.425	.352	.	.471
	LeadQues	.033	.026	.470	.471	.
N	PostTest3difference_Case BC	173	173	173	173	173
	CumAve	173	173	173	173	173
	VirtualSlides	173	173	173	173	173
	InterAnnotate	173	173	173	173	173
	LeadQues	173	173	173	173	173

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.196	4	.049	1.594	.178[b]
	Residual	5.164	168	.031		
	Total	5.360	172			

a. Dependent Variable: PostTest3difference_Case BC

b. Predictors: (Constant), LeadQues, InterAnnotate, VirtualSlides, CumAve

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	.400	.145		2.752	.007					
	CumAve	.146	.178	.063	.820	.413	.036	.063	.062	.974	1.026
	VirtualSlides	.035	.027	.099	1.307	.193	.094	.100	.099	.996	1.004
	InterAnnotate	-.023	.027	-.064	-.851	.396	-.062	-.066	-.064	.999	1.001
	LeadQues	-.051	.026	-.149	-1.952	.053	-.140	-.149	-.148	.978	1.022

a. Dependent Variable: PostTest3difference_Case BC

Note. LeadQues are Guiding Questions

Predictive ability of Learning Interventions Within High-, Mid- and Low-Quintile Student Class Ranks for MCQ Directed to Cases B-C in PostTests 1-3 Analyzed by Multiple Regression

Within High-Quintile Class Rank - PostTest 1- Cases B-C

Descriptive Statistics

	Mean	Std. Deviation	N
PostTest1 difference_Case BC	.7715	.11929	35
VirtualSlides	.4857	.50709	35
InterAnnotate	.4571	.50543	35
LeadQues	.5714	.50210	35

Correlations

		PostTest1 difference_Case BC	VirtualSlides	InterAnnotate	LeadQues
Pearson Correlation	PostTest1 difference_Case BC	1.000	.066	.016	.158
	VirtualSlides	.066	1.000	.026	.033
	InterAnnotate	.016	.026	1.000	.215
	LeadQues	.158	.033	.215	1.000
Sig. (1-tailed)	PostTest1 difference_Case BC	.	.353	.464	.183
	VirtualSlides	.353	.	.441	.425
	InterAnnotate	.464	.441	.	.107
	LeadQues	.183	.425	.107	.
N	PostTest1 difference_Case BC	35	35	35	35
	VirtualSlides	35	35	35	35
	InterAnnotate	35	35	35	35
	LeadQues	35	35	35	35

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.014	3	.005	.309	.819[b]
	Residual	.470	31	.015		
	Total	.484	34			

a. Dependent Variable: PostTest1 difference_Case BC

b. Predictors: (Constant), LeadQues, VirtualSlides, InterAnnotate

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	.745	.040		18.735	.000					
	VirtualSlides	.014	.042	.061	.346	.732	.066	.062	.061	.999	1.001
	InterAnnotate	-.005	.043	-.020	-.113	.911	.016	-.020	-.020	.953	1.049
	LeadQues	.038	.043	.160	.884	.384	.158	.157	.156	.953	1.049

a. Dependent Variable: PostTest1 difference_Case BC

Within Mid-Quintiles Class Rank - PostTest 1- Cases B-C

Descriptive Statistics

	Mean	Std. Deviation	N
PostTest1 difference_Case BC	.6988	.15081	105
VirtualSlides	.50	.502	105
InterAnnotate	.49	.502	105
LeadQues	.52	.502	105

Correlations

		PostTest1 difference_Case BC	VirtualSlides	InterAnnotate	LeadQues
Pearson Correlation	PostTest1 difference_Case BC	1.000	.005	-.129	-.119
	VirtualSlides	.005	1.000	-.048	.067
	InterAnnotate	-.129	-.048	1.000	.011
	LeadQues	-.119	.067	.011	1.000
Sig. (1-tailed)	PostTest1 difference_Case BC	.	.481	.095	.114
	VirtualSlides	.481	.	.314	.248
	InterAnnotate	.095	.314	.	.456
	LeadQues	.114	.248	.456	.
N	PostTest1 difference_Case BC	105	105	105	105
	VirtualSlides	105	105	105	105
	InterAnnotate	105	105	105	105
	LeadQues	105	105	105	105

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.072	3	.024	1.055	.372[b]
	Residual	2.293	101	.023		
	Total	2.365	104			

a. Dependent Variable: PostTest1 difference_Case BC

b. Predictors: (Constant), LeadQues, InterAnnotate, VirtualSlides

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	.735	.029		25.104	.000					
	VirtualSlides	.002	.030	.006	.066	.947	.005	.007	.006	.993	1.007
	InterAnnotate	-.038	.029	-.127	-1.296	.198	-.129	-.128	-.127	.998	1.003
	LeadQues	-.035	.030	-.118	-1.199	.233	-.119	-.118	-.118	.995	1.005

a. Dependent Variable: PostTest1 difference_Case BC

Within Low-Quintile Group - PostTest 1- Cases B-C

Descriptive Statistics

	Mean	Std. Deviation	N
PostTest1 difference_Case BC	.5985	.16465	33
VirtualSlides	.45	.506	33
InterAnnotate	.52	.508	33
LeadQues	.48	.566	33

Correlations

		PostTest1 difference_Case BC	VirtualSlides	InterAnnotate	LeadQues
Pearson Correlation	PostTest1 difference_Case BC	1.000	-.226	-.439	-.067
	VirtualSlides	-.226	1.000	.277	-.248
	InterAnnotate	-.439	.277	1.000	-.244
	LeadQues	-.067	-.248	-.244	1.000
Sig. (1-tailed)	PostTest1 difference_Case BC	.	.103	.005	.355
	VirtualSlides	.103	.	.059	.082
	InterAnnotate	.005	.059	.	.086
	LeadQues	.355	.082	.086	.
N	PostTest1 difference_Case BC	33	33	33	33
	VirtualSlides	33	33	33	33
	InterAnnotate	33	33	33	33
	LeadQues	33	33	33	33

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.214	3	.071	3.165	.039[b]
	Residual	.654	29	.023		
	Total	.868	32			

a. Dependent Variable: PostTest1 difference_Case BC

b. Predictors: (Constant), LeadQues, InterAnnotate, VirtualSlides

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	.727	.054		13.571	.000					
	VirtualSlides	-.051	.056	-.155	-.909	.371	-.226	-.166	-.146	.889	1.125
	InterAnnotate	-.146	.055	-.449	-2.628	.014	-.439	-.439	-.423	.891	1.123
	LeadQues	-.063	.049	-.215	-1.271	.214	-.067	-.230	-.205	.905	1.105

a. Dependent Variable: PostTest1 difference_Case BC